UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS
Departamento de Física de la Tierra, Astronomía y Astrofísica II

ANALYSIS OF SURFACE WIND OVER COMPLEX TERRAIN: A DYNAMICAL DOWNSCALING STUDY WITH THE WRF MODEL.

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Pedro A. Jiménez Muñoz

Bajo la dirección de los doctores

J. Fidel González Rouco
Jorge Navarro Montesinos

Madrid, 2010

- **ISBN: 978-84-693-2391-5**

Analysis of surface wind over complex terrain: a dynamical downscaling study with the WRF model

Tesis doctoral
Pedro A. Jiménez Muñoz
Madrid, Julio 2009

Analysis of surface wind over complex terrain:

a dynamical downscaling study with the WRF model

Análisis del viento superficial en terreno complejo:

simulación regional con el modelo WRF

Memoria que presenta

Pedro Angel Jiménez Muñoz

para optar al grado de

Doctor en Ciencias Físicas

Directores:

Dr. J. Fidel González Rouco
Dr. Jorge Navarro Montesinos

Departamento de Física de la Tierra, Astronomía y Astrofísica II
Facultad de Ciencias Físicas
Universidad Complutense de Madrid

III

Descending from the planetary space and firmament to the surface of our earth, we find a vast variety of phenomena connected with the conversion of living force and heat into one another, which speak in language which cannot be misunderstood of the wisdom and beneficence of the Great Architect of nature. The motion of air which we call wind arises chiefly from the intense heat of the torrid zone compared with the temperature of the temperate and frigid zones. Here we have an instance of heat being converted into the living force of currents of air. These currents of air, in their progress across the sea, lift up its waves and propel the ships; whilst in passing across the land they shake the trees and disturb every blade of grass. The waves by their violent motion, the ships by their passage through a resisting medium, and the trees by the rubbing of their branches together and the friction of their leaves against themselves and the air, each and all of them generate heat equivalent to the diminution of the living force of the air which they occasion. The heat thus restored may again contribute to raise fresh currents of air; and thus the phenomena may be repeated in endless succession and variety.

J. Joule, 1847*

* A lecture at St. Ann's Church Reading-Room, Manchester, 28 April, 1847; reported in the Manchester Courier, 5 and 12 May, 1847; reprinted in *The Scientific Papers of James Prescott Joule*, The Physical Society, London, 1884, pp. 265-76.

Contents

List of acronyms

AR	Autoregressive
AEMET	Agencia Estatal de Meteorología
CA	Cluster analysis
CFN	Comunidad Foral de Navarra
CIEMAT	Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
CLA	Complete linkage algorithm
ECMWF	European Centre for Medium-Range Weather Forecast
EV	Ebro valley subregion
GCM	General circulation model
GN	Gobierno de Navarra
INM	Instituto Nacional de Meteorología
IGRA	Integrated global radiosonde archive
IP	Iberian Peninsula
MM5	The 5^{th} generation of the PSU/NCAR mesoscale model
MS	Mountain stations subregion
MYJ	Mellor-Yamada-Janjic
NCAR	National Center for Atmospheric Research
NS	North-to-South oriented subregion
NV	Northern valleys subregion
PCA	Principal component analysis
PBL	Planetary boundary layer
PP	Pressure pattern
QC	Quality control
RMSE	Root mean squared error
RN	Riegos de Navarra
RSTMS	Reduced spatially and temporally masked simulation
SLP	Sea level pressure
SMS	Spatially masked simulation
STMS	Spatially and temporally masked simulation
UCM	Universidad Complutense de Madrid

USGS U.S. Geophysical survey
YSU Yonsei State University
WP Wind pattern
WRF Weather Research and Forecast modelling system

Summary

The surface circulations play a major role in a variety of phenomena operating at the regional scale such as the dispersion of pollutants over a region, the production and transport of dust due to wind erosion or the wind resource evaluation. An accurate surface wind field estimation is therefore an issue of interest for a wide range of applications. This is specially challenging over complex terrain areas wherein the surface wind field can present a high complexity as a consequence of the strong influence that orography produces over the large scale flows through channeling, forced ascents, blocking, etc.

The present study analyzes the capability of a dynamical downscaling performed with the Weather Research and Forecast (WRF) modeling system to reproduce the surface wind variability and climatology over the Comunidad Foral de Navarra (CFN), a complex terrain region located in the northeast of the Iberian Peninsula (IP). Wind observations acquired during the 1992 to 2005 period at 41 stations are used to represent the surface wind field over the area of study. A quality control is imposed to the original observations in order to remove unrealistic records. The surface wind variability and climatology are analyzed using the quality-controlled observations; the knowledge gained is subsequently used to evaluate the capability of the WRF estimation to reproduce the observed behavior. The simulation was made at the highest possible spatial resolution balancing out the interest in getting a representation of surface features that is as realistic as possible, the simulations of the available parameterizations and computational time. This led to simulate the complete observational period at a horizontal resolution of 2 km. This provides a high horizontal resolution to capture the complexity of the terrain, and a reasonable long simulated period to robustly evaluate the influence of the large scale over the surface flow simulation.

The wind variability over the area is analyzed in a first step using the information from the observational network to identify areas of coherent wind behavior. This classification is accomplished with a twofold objective. In first place, it contributes to understand the surface wind variability over the CFN and secondly, it provides an appropriate framework to evaluate the simulation performance. The regionalization is accomplished by classifying together those observational sites with similar temporal wind variability, which allows to identify the subregions with homogeneous wind behavior. A total of four regions that broadly follow the topographic features of the terrain are identified. The

meridional wind variability is rather similar at the four subregions, the zonal variability being responsible for the differences between the subregions. The spectral analysis of wind variability within each subregion reveals a dominant annual cycle and a varying presence of higher-frequency contributions in the subregions. The valley subregions tend to present more variability at high frequencies than do higher-altitude sites.

Evaluating the model performance at the regional scale provides certain advantages than the traditional evaluation at the sites. The analysis reveals the ability of the WRF simulation to reproduce areas of coherent wind behavior. Based on the good performance shown by the simulation during the evaluation, an analysis is performed to infer the wind variability where and when observations are not available. The inference analysis shows that the reduced number of stations reasonable represent the wind variability over the CFN suggesting some potential improvements.

A further understanding of the wind circulations over the CFN is obtained by analyzing the climatological surface flows over the area. This is accomplished by classifying the observed wind fields with similar structure into groups of typical surface circulation or wind patterns (WPs). The resulting classification is also used to evaluate the WRF simulation performance in reproducing the climatological flows over the area. A total of six WPs are identified. The WPs tend to show a northwest or southeast orientation of the flow. This result shows the strong influence that orography produces over the surface circulations, since the valleys over the region are mainly oriented in the NW-SE direction. The frequency of occurrence of WPs shows an annual cycle. In order to understand the large scale mechanisms that originate the regional circulations, the sea level pressure (SLP) fields over the area are classified into pressure patterns (PPs) and related to the WPs already identified. The relationships found between the PPs and the WPs reveal the influence exerted by the ageostrophic balance over the CFN surface circulations, and the wind intensification produced by the pressure gradient along the valleys.

The WRF simulation reproduces the spatial structure of the climatological flow of the WPs. However, the simulation underestimates the wind on the mountain stations (the windiest sites) and tends to overestimate the wind at the valley subregions where wind speed tends to be lower, therefore causing an underestimation of the spatial variability of the wind speed field. In spite of the high spatial resolution, this behavior is partially related with an still inaccurate representation of some orographic features. The accuracy of the downscaling to reproduce the WPs under their associated PPs is also evaluated to understand the influence of the large scale forcings in the performance of the simulation. For a given WP there is not a PP under which the downscaling clearly outperforms the accuracy displayed by the rest of PPs. Limitations to reproduce the large scale structure in some of the synoptic situations lead to inaccurate downscaled winds. This indicates that the expected improvement obtained with a high spatial resolution in the dynamical downscaling may be at least partially hampered if the large scale cannot be appropriately represented.

Resumen

Entender y predecir la variabilidad del campo de viento a nivel regional es importante para una gran variedad de estudios como pueden ser el transporte y la dispersión de contaminantes sobre una zona, la toma de decisiones en situaciones de riesgo relacionadas con incendios forestales o con situaciones de vientos extremos, o la estimación de la potencia que van a producir los parques eólicos. Esta última aplicación está siendo muy demandada recientemente como consecuencia del gran desarrollo experimentado por la energía eólica en los últimos años (Goswami, 2008; Lu et al., 2009).

Los vientos superficiales medios son el reflejo de la circulación general de la atmósfera, que trata de compensar el calentamiento irregular de la superficie terrestre transportando energía desde el ecuador a los polos (e.g. Lorentz, 1967; Holton, 2004; Schneider, 2006). El viento promedio zonal procede del este en latitudes bajas, del oeste en latitudes medias y del este en latitudes altas. Los vientos promedios meridianos son más débiles y están orientados hacia el norte en regiones de vientos zonales del oeste, y hacia el sur en regiones con vientos superficiales del este. Esta estructura del viento superficial puede reconocerse en los promedios de las circulaciones superficiales del reanálisis de ERA-40 (Simmons and Gibson, 2000; Uppala et al., 2005) mostradas en la Figura 1.1.

Los vientos asociados con la circulación general de la atmósfera son el resultado de la interacción de procesos radiativos y efectos rotacionales que actúan en la atmósfera de una Tierra esférica de superficie uniforme. La heterogeneidad de la superficie terrestre introduce variabilidad adicional en el campo de viento. La influencia de la orografía sobre las circulaciones atmosféricas ha sido foco de numerosos estudios tanto teóricos (Queney, 1948; Scorer, 1949; Jackson and Hunt, 1975; Sykes, 1980; Hunt et al., 1988) como prácticos (Mason and Sykes, 1979; Jenkins et al., 1981; Mason and King, 1985; Taylor and Teunissen, 1987; Vosper et al., 2002). Los efectos orográficos pueden dividirse en forzamientos dinámicos y térmicos asociados con la interación de la orografía con la dinámica y radiación atmosférica, respectivamente (Whiteman, 2000). Los forzamientos dinámicos modifican la circulación atmosférica a gran escala a través de canalizaciones, ascensos forzados, bloqueos, etc. Los efectos térmicos introducen variabilidad adicional, generando circulaciones diarias como los vientos de valle o las circulaciones montaña-valle (Wagner, 1938; Defant, 1949). Los efectos de la orografía sobre las circulaciones atmosféricas a gran escala pueden observarse en la Figura 1.1. El Hemisferio Norte pre-

senta más continentes y por lo tanto la circulación zonal se interrumpe más que en el Hemisferio Sur (Manabe and Terpstra, 1974). La hetereogeneidad de las propiedades físicas de la superficie terrestre originadas como consecuencia del distinto tipo de suelo o la distinta cubierta vegetal pueden inducir circulaciones térmicas adicionales. Un claro ejemplo de esta clase de circulaciones es la brisa marina (Simpson, 1994).

La variabilidad del viento superficial sobre una región está por lo tanto controlada en buena medida por su su posición geográfica, que determina la influencia de la circulación general, y por las carácterísticas del terreno en la región. Su estudio puede abordarse desde un punto de vista observacional, que proporciona información de la variabilidad del viento (e. g. Baker et al., 1978; Martner and Marwitz, 1982; Wendland, 1982; Klink, 2002; Archer and Jacobson, 2003, 2004, 2005; McVicar et al., 2008), o, alternativamente, se pueden usar modelos numéricos que permiten una mayor comprensión de los procesos físicos involucrados en el flujo regional, sus interacciones con las carácterísticas del terreno, así como otros factores a mesoscala y a gran escala (e. g. Mahrer and Pielke, 1977; Mahrer et al., 1985; Rife et al., 2004).

Los estudios basados en observaciones intentan proporcionar una interpretación lo más precisa posible de la variabilidad del viento a escala regional asumiendo que las observaciones *in situ* son representativas de una cierta área en su entorno y que la integración de la información de todos los emplazamientos es suficiente para proporcionar una buena cobertura espacial a nivel regional. La fiabilidad de la representación del flujo regional puede estar por lo tanto limitada por la calidad y disponibilidad de datos observacionales. La topografía influye en la variabilidad del flujo como consecuencia de los efectos dinámicos y térmicos generados por el terreno (Whiteman, 2000), de tal forma que un campo de viento más complejo requiere una red de observaciones más densa para capturar mejor la variabilidad espacial de las circulaciones superficiales.

Los estudios basados en modelos numéricos emplean las ecuaciones de la hidrodinámica (e.g. Cushman-Roisin, 1994; Dutton, 2002; Holton, 2004) para proporcionar estimaciones físicamente consistentes del estado atmosférico. Al tratarse de un sistema no lineal de ecuaciones en derivadas parciales no es posible obtener una solución analítica y se necesita encontrar una solución aproximada (Bjerknes, 1904). La solución es una función de las condiciones iniciales que necesitan ser facilitadas de antemano.

Los algoritmos numéricos usados para simular la evolución de la atmósfera a nivel global se conocen como modelos de circulación general (MCGs). Los MCGs son capaces de reproducir la alternancia del este y oeste de los vientos superficiales, y han contribuido a comprender mejor la circulación general de la atmósfera (Phillips, 1956; Lewis, 1998). No obstante, la integración de las ecuaciones de la hidrodinámica sobre toda la atmósfera requiere recursos computacionales elevados que limitan severamente la discretización espacial empleada. La resolución horizontal típica de los MCGs es del orden de un grado de latitud x longitud (111 km de brazo de red en el Ecuador). Por tanto, los MCGs pueden producir estimaciones adecuadas en regiones de terreno homogéno, pero sus estimaciones deben de interpretarse con cuidado a medida que la heterogeneidad de la superficie terrestre aumenta. Este es el caso de las regiones de terreno complejo, en donde el campo de viento superficial puede presentar una alta variabilidad espacial como consecuencia de las modificaciones que la orografía produce sobre los flujos a escala sinóptica.

Teóricamente, esta limitación de los MCGs podría ser evitada incrementando la resolución horizontal para poder capturar los detalles a nivel regional. Sin embargo, el coste computacional necesario resulta prohibitivo. Por lo tanto, para obtener información a nivel regional es necesario emplear procedimientos alternativos a los MCGs. Con este propósito, y basándose en el buen funcionamiento de los MCGs para reproducir los campos de la escala sinóptica, se desarrollaron técnicas para aumentar la resolución espacial o técnicas de *"downscaling"* (Giorgi and Means, 1991; von Storch, 1995; Hewitson and Crane, 1996; Wilby and Wigley, 1997). Estas técnicas pueden dividirse en dos grandes grupos: estadísticas y dinámicas, dependiendo de los procedimientos empleados para estimar la información a escala regional.

Las técnicas estadísticas establecen relaciones empíricas entre los campos a gran escala y las observaciones existentes durante un periodo de calibración (Zorita et al., 1992; von Storch et al., 1993; Wilby et al., 1998; González-Rouco et al., 2000; Xoplaki et al., 2004). Una vez establecida, la relación empírica puede usarse para estimar la variabilidad regional en periodos en los que no hay observaciones disponibles utilizando las salidas de un MCG como campos predictores. Por lo tanto, se está asumiendo implícitamente que la relación encontrada durante el periodo de calibración se mantiene en otros periodos temporales. Esta hipótesis no puede garantizarse y constituye una limitación. Además, las técnicas estadísticas sólo pueden proporcionar estimaciones en aquellos lugares donde hay observaciones disponibles. Por otro lado, presentan la ventaja de requerir pocos recursos computacionales para obtener la información a escala regional una vez que la relación con la gran escala ha sido encontrada. La técnicas estadísticas han sido ampliamente utilizadas con variables relacionadas con la precipitación y la temperatura del aire en superficie (e.g. Kim et al., 1984; Wigley et al., 1990; Noguer, 1994; González-Rouco et al., 2000; Huth, 2002; Xoplaki et al., 2003a,b, 2004; Vrac et al., 2007; Paul et al., 2008; Busuioc et al., 2008; Li and Smith, 2009). Por el contrario, un análisis exhaustivo de su capacidad para estimar variables relacionadas con el viento ha sido poco explorado. Un estudio pionero aplicó metodologías lineales basadas en el análisis de correlación canónica para reconstruir parámetros asociados con el viento durante el último siglo en varios emplazamientos en el Atlántico Norte (Kaas et al., 1996). Faucher et al. (1999) usaron técnicas estadísticas para reconstruir el viento observado en boyas situadas al oeste de Canadá. Otros estudios realizan estimaciones de diferentes variables meteorológicas entre las que se encuentran algunas relacionadas con el viento (e.g. Gutiérrez et al., 2004; Cheng et al., 2007). El área donde mayor uso han tenido este tipo de técnicas ha sido la estimación de los cambios de la velocidad del viento en escenarios de futuro asociados con un potencial cambio climático (Bogardi and Matyasovszky, 1996; Sailor et al., 2000; Pryor et al., 2005b,c, 2006; Najac et al., 2009).

La simulaciones regionales se basan en principios físicos muy similares a aquellos usados por los MCGs para aumentar la resolución espacial (Krishnamurti and Bounoua, 1995; Pielke, 2002). Sin embargo, la evolución atmosférica se simula sólo sobre una región de la superficie terrestre. Esto permite aumentar la resolución espacial sin comprometer los recursos computacionales, alcanzando resoluciones horizontales típicas que varían entre unas decenas de km hasta 1 km. El aumento de la resolución espacial proporciona una mejor representación de la heterogeneidad de la superficie terrestre que debería traducirse en una mejor estimación de la variabilidad regional (Mass et al., 2002). El

incremento de la resolución espacial también permite mejorar la representación de los procesos físicos que se parametrizan en los volúmenes en los cuales se ha discretizado la atmósfera por tener una escala de resolución inferior a la del modelo (Stensrud, 2007). La mejora en la representación de los procesos físicos en los modelos regionales también debería contribuir a aumentar su realismo con respecto a las simulaciones proporcionadas por los MCGs.

Las simulaciones regionales se realizan con modelos de área limitada o modelos mesoscalares (Black, 1994; Grell et al., 1994; Bubnova et al., 1995; Cotton et al., 2003; Skamarock et al., 2005). Estos modelos usan las salidas de los MCGs para obtener las condiciones iniciales y de contorno necesarias para aproximar una solución. Al principio, los modelos de área limitada se usaron para realizar predicciones numéricas del tiempo o para analizar fenómenos a mesoscala (Physick, 1988; Warner, 1989). La idea de usar los modelos de área limitada para estudiar los impactos regionales del sistema climático global fue propuesta por Dickinson et al. (1989). Tras dicha propuesta, no se tardó mucho en realizar las primeras simulaciones regionales de un mes de duración (e.g. Giorgi and Bates, 1989; Giorgi, 1990; Marinucci et al., 1995; Walsh and McGregor, 1995); estos experimentos se irían aumentando en longitud temporal y resolución espacial con la mayor disponibilidad de medios computacionales (e.g. Giorgi et al., 1993; Christensen et al., 1998). El campo experimentó un desarrollo enorme en los años sucesivos (McGregor, 1997; Wang et al., 2004; Giorgi, 2006). Algunos estudios han evaluado la capacidad de simulaciones regionales relativamente largas en reproducir el campo de viento en superficie (Buckley, 2004; Conil and Hall, 2006; Walter et al., 2006; Kanamitsu and Kanamaru, 2007). Sin embargo, la realización de simulaciones de varios años a muy alta resolución espacial (unos pocos km) todavía resulta computacionalmente costoso y esta vía no ha sido aún muy explorada a pesar de su potencial para analizar la variabilidad espaciotemporal del viento en regiones de terreno complejo. Normalmente se suele simular un periodo temporal corto, de unos pocos meses o la duración de una campaña observacional, para dimensionar los recursos computacionales necesarios (e.g. Georgelin and Richard, 1996; Masson and Bougeault, 1996; Zhong and Fast, 2003; Cairns and Corey, 2003; Rife et al., 2004; Rife and Davis, 2005; Zagar et al., 2006; Soriano et al., 2006; Pérez-Landa et al., 2007). Estos estudios han contribuido a entender mejor la influencia que la orografía ejerce sobre las circulaciones superficiales o a evaluar la bondad de una simulación bajo determinados forzamientos sinópticos. Sería interesante explorar la capacidad de reproducir la variabilidad de baja frecuencia de las circulaciones en superficie por medio de simulaciones regionales realizadas a alta resolución espacial. Otro aspecto interesante a analizar consistiría en evaluar la mejora introducida en las estimaciones como consecuencia de aumentar de resolución horizontal.

Por lo tanto, necesitamos evaluar mejor la capacidad de las técnicas de aumento de resolución espacial, tanto las estadísticas como las basadas en simulaciones regionales, antes de usarlas de forma rutinaria para estimar las circulaciones superficiales.

Objetivos

El objetivo principal de este trabajo es analizar la capacidad del modelo mesoscalar para la investigación y predicción del tiempo (WRF, Skamarock et al., 2005) de reproducir la variabilidad y la climatología de las circulaciones superficiales en una zona de terreno complejo. Para ello se ha realizado una simulación en alta resolución espacial que cubre un intervalo temporal de trece años. El análisis se lleva a cabo con medias diarias del viento. Esta decisión filtra parcialmente los efectos asociados con forzamientos térmicos, y permite centrar el estudio en la influencia de los forzamientos dinámicos asociados a la interación de la dinámica atmosférica a escala sinóptica con la orografía regional.

Para proporcionar una evaluación adecuada es necesario entender el comportamiento del viento sobre la zona. Por lo tanto, este trabajo analiza primero la variabilidad y la climatología del viento en superficie con observaciones *in situ* para seguidamente analizar la habilidad de la simulación numérica en reproducir el comportamiento del viento observado. Con este propósito, los análisis realizados con las observaciones están orientados a entender las circulaciones superficiales sobre la zona de estudio y a la vez proporcionar un marco adecuado para la evaluación. Por tanto, hay dos objetivos parciales en la Tesis: i) analizar la variabilidad y climatología del viento superficial en una zona de terreno complejo y ii) evaluar la capacidad de reproducir el comportamiento observado mediante una simulación regional realizada con el modelo WRF.

El área de estudio seleccionada es la Comunidad Foral de Navarra (CFN), una región de terreno complejo situada en el noreste de la Península Ibérica (IP, Fig. 1.2). La CFN se caracteriza por sus altas velocidades de viento que han favorecido la instalación de numerosos parques eólicos sobre la zona en los últimos años (Fairless, 2007; García-Bustamante et al., 2008, 2009). La CFN tiene una red de estaciones meteorológicas razonablemente densa con observaciones de viento disponibles desde 1992 lo que ha permitido llevar a cabo varios análisis del viento en la zona (García et al., 1998; Torres et al., 1999, 2005; Jiménez et al., 2008, 2009b). Dada la existencia de una orografía relativamente compleja y una buena red de observaciones, la CFN constituye un área interesante para evaluar la capacidad de reproducir el viento en superficie mediante simulaciones regionales.

Aproximación conceptual

No hay ninguna razón obvia para justificar el uso de una simulación regional en detrimento de técnicas estadísticas, porque ambas metodologías presentan sus ventajas y desventajas y sería interesante evaluar el comportamiento de ambas. Entre las ventajas que presenta la simulación regional sobre las técnicas estadísticas cabe destacar su mayor cobertura espacial en tres dimensiones, su consistencia física y su relativa independencia de las observaciones.

La comparación de las simulaciones numéricas con observaciones se suele denominar proceso de “verificación” o “validación”. Sin embargo, se puede argumentar que la verificación o validación de los modelos numéricos que describen los sistemas naturales es imposible (Oreskes et al., 1994; Oreskes, 1998). Las simulaciones pueden compararse con

las observaciones para "confirmar" que la simulación es adecuada para reproducir ciertos aspectos de la realidad en un caso particular, pero nada puede asegurarnos que la simulación será adecuada en otras situaciones. Cuanto mayor sea el número de análisis en los que la simulación muestre resultados adecuados, mayor será la posibilidad de que la simulación proporcione resultados aceptables en otras situaciones comparables. En este sentido, las simulaciones regionales sólo pueden evaluarse en términos relativos (Oreskes et al., 1994).

Por otro lado, la comparación de las simulaciones con información local es siempre delicada (von Storch, 1995). Normalmente, las observaciones *in situ* se comparan con la simulación en el punto de la rejilla computacional más cercano a la estación de medida (Cox et al., 1998; Hanna and Yang, 2001; Buckley, 2004). Sin embargo, merece la pena destacar dos razones por las que este tipo de comparaciones resultan problemáticas. Por un lado, las simulaciones proporcionan variables promediadas en volúmenes con propiedades homogéneas debido a la discretización espacial y al promedio de Reynolds que se aplica para simplificar las ecuaciones (Reynolds, 1895). Por tanto, su comparación con observaciones *in situ* es discutible en aquellos emplazamientos sometidos a efectos locales no considerados en el volumen simulado. Por otro lado, la discretización introducida suaviza la complejidad del terreno y las propiedades físicas de la superficie terrestre. Esto puede conducir a una situación en la que la selección del volumen que incluye el emplazamiento de la observación no es la más apropiada para representar la variabilidad observada; siendo la simulación en volúmenes cercanos más adecuada. Nos referiremos a este último problema como el del *error de representación*. Existen diferentes metodologías para mitigar estos efectos adversos. Por ejemplo, se pueden emplear modelos de diagnóstico para aumentar la resolución espacial proporcionada por las simulaciones regionales (e.g. Chandrasekar et al., 2003; Ludwig et al., 2006), o comparar los campos espaciales suavizados de funciones ortogonales empíricas (EOFs, del inglés *"empirical orthogonal functions"*) procedentes de un análisis de componentes principales (PCA, del inglés *"principal component analysis"*) aplicado a la simulación y a las observaciones (Volmer et al., 1984). Un procedimiento interesante fue propuesto por Reid y Turner (2001) que compararon los promedios en los volúmenes procedentes de una simulación regional realizada a baja resolución horizontal (40 km) con promedios de observaciones en regiones definidas de forma subjetiva. El promedio de las series observacionales necesario para obtener las series regionales filtra la variabilidad local (ruido), reforzando la señal regional. Esto condujo a mejorar la comparación entre las salidas del modelo y las observaciones respecto al método más tradicional de comparar con el punto de la rejilla más cercano.

En esta Tesis se usa un método automático para identificar las regiones con similar variabilidad temporal del viento sin ningún conocimiento *a priori* del comportamiento del viento sobre la zona. En este sentido se usa un procedimiento más objetivo que el de Reid y Turner (2001). La regionalización permite analizar la variabilidad espacial y temporal del viento sobre la zona de estudio (Capítulo 4). A su vez, la identificación de áreas en las que el viento se comporta de forma homogénea también proporciona un marco adecuado para evaluar la capacidad de la simulación regional para reproducir la variabilidad superficial del campo de viento (Capítulo 5). En concreto, el promedio de las series temporales necesario para obtener la serie regional mitiga los efectos locales en las observaciones y los *errores de representación* aleatorios en la simulación, proporcionando series temporales

más adecuadas para su comparación. Basándose en los resultados obtenidos durante la evaluación, la mayor cobertura espacial y temporal que proporciona la simulación regional se usa para analizar la variabilidad del viento en lugares y periodos en los que no hay observaciones disponibles (Capítulo 5).

El estudio de la climatología de las circulaciones superficiales proporciona información complementaria a los estudios de variabilidad del campo de viento. Las circulaciones climatológicas sobre la zona se estudian mediante la identificación de patrones de viento que caracterizan las circulaciones locales (Capítulo 6). Con este propósito, los campos de viento observados se clasifican en grupos con similar tipo de circulación, lo que permite identificar las circulaciones predominantes sobre la zona y su relación con el clima regional. Los campos de presión a nivel del mar también se clasifican con el fin de identificar los patrones sinópticos típicos y poder relacionarlos con las circulaciones de viento y así entender los mecanismos responsables de la variabilidad atmosférica regional. Este análisis permite también analizar la capacidad de la simulación regional para reproducir la climatología de las circulaciones sobre la región (Capítulo 7). Cada tipo de circulación de viento agrupa un gran número de días lo que contribuye a aumentar la confianza al generalizar los resultados obtenidos en la evaluación (ver la discusión previa sobre Oreskes et al., 1994 y Oreskes, 1998). Además, el conocimiento de las situaciones a gran escala que pueden producir cada tipo de circulación en superficie permite evaluar la capacidad del modelo en reproducir tales circulaciones bajo distintos forzamientos sinópticos.

Aportaciones fundamentales

A continuación se describen los objetivos parciales de cada sección de la Tesis así como los resultados más relevantes.

- **Datos observacionales**

  En primera instancia, los datos observacionales se sometieron a un proceso de depuración con el objetivo de suprimir o corregir medidas potencialmente erróneas (Capítulo 2). Los datos observacionales consisten en promedios de 10 o 30 minutos del módulo y la dirección del viento registrados desde 1992 hasta 2005 en 41 estaciones meteorológicas automáticas (Fig. 1.2). Existen numerosas referencias para la aplicación de controles de calidad y homogenización de variables relacionadas con la temperatura y la precipitación, sin embargo, hay muy pocos estudios que se centren en variables relacionadas con el viento (DeGaetano, 1997, 1998; Graybeal, 2006). En particular, no hay ningún trabajo previo cuyo objetivo sea mejorar la calidad de observaciones de alta frecuencia del viento registradas en estaciones meteorológicas automáticas como es el caso de los datos empleados en esta Tesis.
  Los datos de viento y dirección son sometidos a un control de calidad (Fig. 2.1) para suprimir las medidas cuestionables y, seguidamente, se suprimen o corrigen los periodos con diferencia de medias o varianza en las distintas series temporales. El proceso de depuración de la calidad de los datos descrito en este trabajo no solo contribuye con la creación de una base de datos de viento de mayor calidad, sino que

también permite entender un poco mejor la problemática relacionada con la medida de esta variable.
Los resultados más relevantes de esta Sección están siendo redactados para su publicación (Jiménez et al., 2009d).

- **La simulación regional**

  La simulación regional se realiza con el modelo WRF (Skamarock et al., 2005) y comprende todo el periodo observacional disponible, desde 1992 hasta 2005, con una alta resolución espacial de 2 km sobre la CFN (Capítulo 3). Estas carácterísticas de la simulación permiten representar de forma realista la orografía compleja de la CFN y a la vez proporcionan un periodo suficientemente largo como para poder analizar la variabilidad de baja frecuencia del viento. El valor de la simulación debe de contemplarse teniendo en cuenta el alto coste computacional necesario que demandan los modelos mesoscalares. No se tiene constancia de ningún estudio previo que evalúe la capacidad de una simulación regional llevada a cabo a tan alta resolución espacial y sobre un periodo temporal tan extenso como es el caso descrito en esta Tesis.

- **Variabilidad del campo de viento**

  La variabilidad del viento sobre la CFN se estudia mediante la identificación de aquellos emplazamientos con variabilidad temporal similar del viento. Esta regionalización permite determinar las zonas con comportamiento homogéneo del viento. El objetivo que se persigue con esta clasificación es doble. Por un lado permite estudiar la variabilidad del viento sobre la zona, mientras que por otro lado permite realizar la evaluación de la simulación por subregiones en vez de por emplazamientos.
  La regionalización se lleva a cabo por medio de dos metodologías basadas en PCAs. Esto permite comparar resultados y así determinar las subregiones de forma más robusta. El primer método de regionalización realiza un análisis de conglomerados (del inglés *"cluster analysis"*) de las componentes de los autovectores procedentes del PCA (Romero et al., 1999b). El segundo método rota los autovectores para identificar las regiones de comportamiento homogéneo del viento (White et al., 1991). Aunque este tipo de clasificaciones se emplean con relativa frecuencia en la regionalización de variables escalares (e.g. Dyer, 1975; White et al., 1991; Coronato and Bisigato, 1998; Bonell and Sumner, 1992; Fovell and Fovell, 1993; Sanchez-Lorenzo et al., 2007, 2008; Burlando et al., 2008) no hay constancia de que se hayan aplicado al viento horizontal, una variable vectorial, como es el caso de esta Tesis.
  Los regiones identificadas por el método de conglomerados y el método basado en la rotación pueden observarse en las Figuras 4.8c y 4.11a, respectivamente. Ambas producen regiones equivalentes distribuidas de acuerdo con la orografía de la zona. Hay una primera subregión que engloba estaciones en el valle del Ebro (círculos en Fig. 4.11a). Otra subregión engloba las estaciones localizadas en la cima de montañas (cuadrados en Fig. 4.11a). La tercera subregión esta formada por un conjunto de estaciones con una orientación Norte-Sur en el centro de la CFN (rombos en Fig. 4.11a). La última subregión está fundamentalmente formada por estaciones en valles

al norte del valle del Ebro (triángulos en Fig. 4.11a). La principal diferencia entre las dos clasificaciones es que la que se basa en el análisis de conglomerados genera una zona con regiones pequeñas que son agrupadas en la subregión con orientación Norte-Sur por el método basado en la rotación.
La variabilidad meridiana del viento es similar en todas las regiones, siendo responsable de su diferente variabilidad la componente zonal del viento. El análisis espectral de la variabilidad del viento en las subregiones revela la presencia de una onda anual dominante y diferentes contribuciones de alta frecuencia en las distintas regiones (Fig. 4.15). Las regiones de valle tienden a mostrar mayor actividad a altas frecuencias que la subregión de montaña.
Los resultados más importantes de esta Sección de la Tesis han sido publicados como Jiménez et al. (2008).

- **Evaluación de la simulación regional: variabilidad**

  El objetivo principal de esta sección es evaluar la habilidad de la simulación regional para reproducir la variabilidad del campo del viento en superficie. El análisis se basa en la regionalización obtenida con las observaciones. Aquí se muestra como la evaluación a nivel regional proporciona mejor concordancia entre observaciones y simulaciones que la tradicional comparación a escala local (Fig. 5.2). Esto se debe a que el promediado de las series observadas/simuladas necesario para obtener las series regionales filtra los efectos locales y *errores de representación* de las series individuales (ruido) y a la vez refuerza la señal regional. El resultado de la evaluación muestra el realismo de la simulación al reproducir las áreas de similar variabilidad del viento. En las subregiones del valle del Ebro y la de montaña la simulación proporciona las mejores estimaciones.
  La adecuada reprodución de la variabilidad espaciotemporal del viento proporcionó confianza para emplear la mayor cobertura espacial y temporal que proporciona la simulación para analizar la variabilidad del viento en periodos y áreas con carencia de observaciones. Este análisis de inferencia sugiere que el reducido número de estaciones es adecuado para representar la variabilidad espacial del viento, aunque convendría aumentar la densidad de estaciones en algunas zonas. Este tipo de estudio puede ser de cierta utilidad a la hora de mejorar el diseño de redes de observaciones meteorológicas.
  Los resultados más importantes de esta Sección de la Tesis están actualmente en proceso de revisión para su publicación (Jiménez et al., 2009a).

- **Climatología del campo de viento**

  Las circulaciones climatológicas sobre la CFN se analizan clasificando juntos aquellos días con similar estructura en el campo de viento. De nuevo el estudio presenta el doble objetivo de entender el comportamiento del viento y a la vez proporcionar un entorno adecuado para la evaluación de la simulación. La clasificación se realizó con dos metodologías diferentes. El primer método clasifica los campos de viento en base a su similitud espacial (Kaufmann and Whiteman, 1999) mientras que el segundo se basa en la variabilidad temporal (Green et al., 1993). Ambas clasificaciones propor-

cionan resultados similares lo que confiere robustez a los seis patrones superficiales de viento identificados (Fig. 6.7). Cinco de estos patrones presentan flujo del noroeste-sureste (Fig. 6.7a-e) lo que indica la gran influencia que ejerce la orografía sobre las circulaciones superficiales, dado que los valles en la CFN se alinean principalmente en esa dirección. La frecuencia de aparición de estos patrones a lo largo del año muestra una marcada estacionalidad (Fig. 6.14).
Los estados atmosféricos asociados a cada uno de los patrones de viento representados por la temperatura, humedad relativa, radiación global y precipitación revelan distintos regímenes de advección dependiendo del patrón de circulación superficial.
Otro de los objetivos parciales en esta sección es analizar los patrones sinópticos responsables de las circulaciones superficiales. Con este fin, los campos de presión a nivel del mar con una estructura similar se clasificaron juntos en lo que se denominaron patrones de presión. Se identificaron ocho patrones de presión (Fig. 6.18). Estos patrones son consistentes con los patrones identificados en otros estudios previos (Soler, 1977; Guardans and Palomino, 1995; Zhang et al., 1997; Trigo and DaCamara, 2000). Los patrones muestran importantes asociaciones con los seis patrones típicos de circulación superficial (Tabla 6.2). Del resultado de analizar estas relaciones se deduce la importancia del balance ageostrófico y la intensifiación del viento ejercida por el gradiente de presión a lo largo del valle del Ebro.

Los resultados más importantes de esta Sección de la Tesis han sido publicados como Jiménez et al. (2009b).

- **Evaluación de la simulación regional: climatología**

El objetivo de esta sección es determinar la habilidad de la simulación para reproducir la climatología de las circulaciones superficiales. El análisis presta especial atención a la influencia de la orografía y las circulaciones a gran escala a la hora de estimar el campo de viento.
En general, la simulación muestra una tendencia a sobrestimar el viento en los valles, las zonas menos ventosas, y a infraestimarlo en las estaciones de montaña, la más ventosas (Fig. 7.1). Esto conlleva a que la simulación infraestime la variabilidad espacial del campo de viento, y está relacionado con el suavizado de la topografía que se aplica para poder realizar la simulación regional. Las estimaciones de la dirección son mejores en la región de montaña que en las regiones de valles. Esto probablemente está asociado con la menor influencia que ejerce la orografía circundante en la región de montaña que está más expuesta a los vientos geostróficos.
La simulación reproduce de forma razonable la estructura del flujo superficial de cada uno de los patrones típicos de viento (Fig. 7.4). Las principales discrepancias están asociadas a limitaciones de la simulación para reproducir ciertas carácterísticas de algunos patrones de presión. En concreto, la simulación regional tiende a infraestimar la presión a nivel del mar al noroeste de la Península Ibérica, lo que causa una infraestimación del gradiente de presión a lo largo de los valles de la CFN, que a su vez se traduce en una infraestimación de la velocidad del viento en los valles en ciertos patrones de viento. En algunos casos esta infraestimación del gradiente de

presión conduce a la simulación de vientos de sentido opuesto al observado. Por tanto, limitaciones en la representación de la topografía y limitaciones para reproducir la gran escala introducen errores en las estimaciones de las circulaciones superficiales observadas.
Los resultados más importantes de esta Sección de la Tesis están siendo redactados para su publicación (Jiménez et al., 2009c).

Conclusiones más relevantes

Este trabajo ha contribuido a comprender mejor la capacidad de una simulación regional de alta resolución espacial para reproducir las circulaciones superficiales en una zona de terreno complejo.

La variabilidad de la componente meridiana del viento es similar en toda la CFN, siendo la variabilidad de la componente zonal la que marca la diferencia en el comportamiento de viento. El viento muestra un ciclo anual dominante con contribuciones de alta frecuencia en los valles.

Las circulaciones predominantes son del noroeste seguidas por las del sureste. Esta es la dirección de los valles más importantes de la CFN, lo que revela la gran influencia que ejerce la orografía sobre las circulaciones superficiales. El balance ageotrófico y la intensificación ejercida por el gradiente de presión a lo largo de los valles de la CFN controlan en buena medida las circulaciones superficiales.

La simulación muestra estimaciones aceptables de la variabilidad del viento en superficie (Fig. 8.1). Se han puesto de manifiesto las ventajas que la evaluación a nivel regional proporciona en comparación con la evaluación a nivel local. Los resultados animan a extender la simulación regional a periodos anteriores a 1992 para poder analizar la variabilidad del viento cuando las observaciones son muy limitadas.

Limitaciones en la representación de la orografía así como problemas para reproducir las circulaciones a escala sinóptica introducen errores en las estimaciones de las circulaciones superficiales. Por tanto, la mejora que cabría esperar como resultado de aumentar la resolución espacial de la simulación regional está limitada por la capacidad del modelo en reproducir las circulaciones a escala sinóptica. Extender la simulación a periodos anteriores a 1992 para analizar la climatología del viento superficial requeriría aplicar alguna clase de corrección. Una posibilidad consistiría en corregir los errores introducidos por cada uno de los ocho patrones de presión identificados sobre la Península Ibérica (Fig. 8.2). Esta simulación corregida podría usarse en el futuro para evaluar la sostenibilidad del gran número de parques eólicos instalados en la CFN. Otra posibilidad para mejorar las estimaciones podría consistir en aplicar técnicas que impongan la estructura de la gran escala (del inglés *"nudging techniques"*) a la simulación regional.

Publicaciones relacionadas con la Tesis en las que ha participado el autor

- García-Bustamante, E., J. F. González-Rouco, P. A. Jiménez, J. Navarro, and J. P. Montávez, 2008: The influence of the Weibull assumption in monthly wind energy estimation. Wind Energy, 11, 483-502.
- García-Bustamante, E., J. F. González-Rouco, P. A. Jiménez, J. Navarro, and J. P. Montávez, 2009: A comparison of methodologies for monthly wind energy estimations. Wind Energy, (en prensa).
- Jiménez, P. A., J. F. González-Rouco, E. García-Bustamante, J. Navarro, J. P. Montvez, J. Vilà-Guerau de Arellano, J. Dudhia, and A. Roldan, 2009a: Surface wind regionalization over complex terrain: evaluation and analysis of a high resolution WRF numerical simulation. J. Appl. Meteor. & Climatol., (en prensa).
- Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, E. García-Bustamante, and J. Navarro, 2009b: Climatology of wind patterns in the northeast of the Iberian Peninsula. Int. J. Climatol., 29, 501-525.
- Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, E. García-Bustamante, J. Navarro, J. Vilá-Guerau de Arellano, J. Dudhia, and A. Roldan, 2009c: High-resolution dynamical downscaling of daily mean surface circulations over complex terrain, (en preparación).
- Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, J. Navarro, E. García-Bustamante, and F. Valero, 2008: Surface wind regionalization in complex terrain. J. Appl. Meteor. & Climatol., 47, 308-325.
- Jiménez, P. A., J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante, 2009d: Quality control and bias correction of high reesolution surface wind observations from automated weather stations, (en preparación).
- Jiménez, P. A., J. P. Montávez, E. García-Bustamante, J. Navarro, J. M. Jiménez-Gutiérrez, E. E. Lucio-Eceiza, and J. F. González-Rouco, 2009e: Diurnal surface wind variations over complex terrain. Física de la Tierra, (en revisión).

1

Introduction

1.1 A challengeable problem

Understanding and forecasting regional wind variability is relevant for a wide variety of phenomena, for example, the transport and dispersion of pollutants along an area, planning and decision making within situations of risk assessment related to the occurrence of extreme events (e.g. forest fires, structural damages in buildings, etc) and power forecasting in wind farms. The latter is becoming an issue of increasing demand to establish national policies in view of the recent developments of renewable energy technologies (Goswami, 2008; Lu et al., 2009).

The mean surface winds are a result of the general circulation of the atmosphere, which tries to compensate the differential heating of the Earth's surface by transporting energy from the equator to the poles (e. g. Lorentz, 1967; Holton, 2004; Schneider, 2006). Mean zonal winds are easterly in low latitudes, westerly in middle latitudes and weak easterly winds are present at high latitudes. The mean meridional winds are weaker and display a poleward (equatorward) component in regions of surface westerlies (easterlies). This structure can be recognized in the averaged surface circulations from the ERA-40 reanalysis project (Simmons and Gibson, 2000; Uppala et al., 2005) displayed in Figure 1.1. The surface zonal/meridional wind alternation is the result of meridional circulation cells within the troposphere (Palmén and Vuorela, 1963; Vuorela and Tuominen, 1964).

The winds associated with the general circulation of the atmosphere are a consequence of radiative processes and rotational effects acting on the atmosphere of a spherical Earth. Considering a nonuniform earth surface introduces additional variability. The topography influence on the atmospheric circulations has been the focus of numerous theoretical (Queney, 1948; Scorer, 1949; Jackson and Hunt, 1975; Sykes, 1980; Hunt et al., 1988) and observational research (Mason and Sykes, 1979; Jenkins et al., 1981; Mason and King, 1985; Taylor and Teunissen, 1987; Vosper et al., 2002). In general terms, the orography produces dynamical and thermal forcings generated by the terrain interaction with the atmospheric dynamics and radiation (Whiteman, 2000). The dynamical forcing modifies the large scale flow due to channeling, forced ascents, blocking, etc. Thermal effects introduce additional wind variability by generating local diurnal circulations such as valley winds or mountain-plain circulations (Wagner, 1938; Defant, 1949). The effects

of topography over the planetary-scale flow can be noticed in Figure 1.1. The northern hemisphere presents more continents and therefore more mountainous regions, and, as a consequence, it shows a less clear zonal circulation than the southern hemisphere where the flow encounters less obstacles (Manabe and Terpstra, 1974). The heterogeneity of the surface physical properties such as the soil type or vegetation cover can potentially generate additional thermally driven circulations. Clear examples of this kind of induced flows are the land/sea breezes (Simpson, 1994).

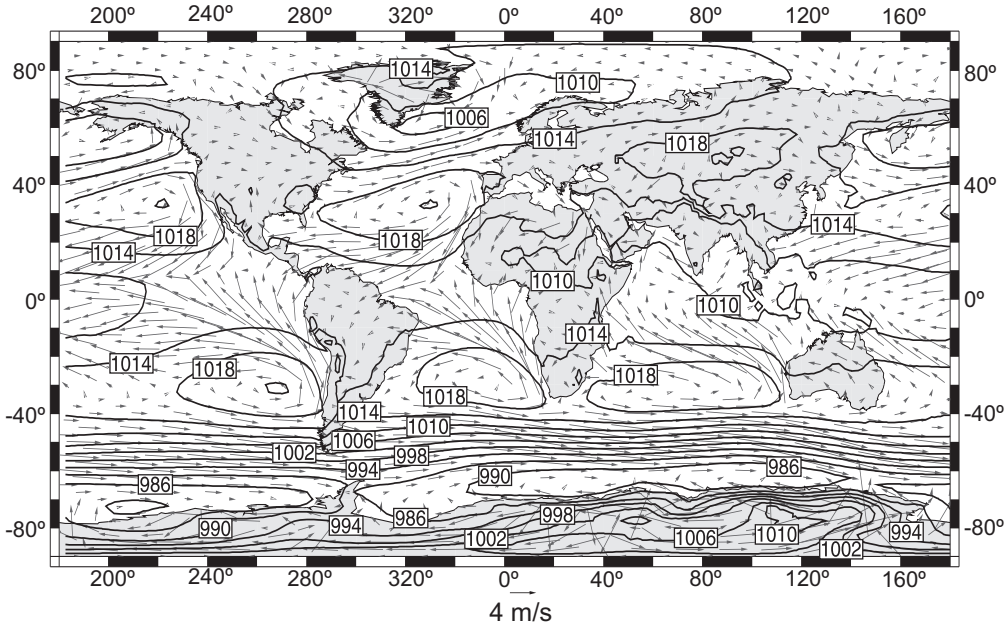

Fig. 1.1: Mean sea level pressure (contour lines) and mean wind (vectors) in the period 1957 - 2002 from the ERA-40 reanalysis project.

The surface wind characteristics over a particular region are therefore largely controlled by its specific location within the globe, which determines the influence of the large scale circulation and by the terrain features. Understanding of the regional wind field can be undertaken from a purely observational standpoint that provides information on the statistics of wind variability (e. g. Baker et al., 1978; Martner and Marwitz, 1982; Wendland, 1982; Klink, 2002; Archer and Jacobson, 2003, 2004, 2005; McVicar et al., 2008). Additionally, using model based approaches allow for a more complete understanding of the physical processes and mechanisms operating on the regional flow, its interactions with landscape features and other driving factors at meso- and large- scales (e. g. Mahrer and Pielke, 1977; Mahrer et al., 1985; Rife et al., 2004).

Observational based studies attempt to address an accurate representation of wind variability at the regional level assuming that local *in situ* information is representative for some area around the site, and that the integration of this information for all sites is sufficient to provide a good regional scale coverage. The reliability on the representation of the regional flow can be however hampered by the quality and availability of wind records. Topography also impinges on the regional representation by increasing the complexity of the flow as a consequence of the dynamical and thermal forcing generated by the orography (Whiteman, 2000). Thus, a more complex wind field requires a denser network of stations to better capture the spatial variability and details of the flow.

Model based assessments lean on the hydrodynamic equations (e.g. Dutton, 2002; Holton, 2004) to provide physically consistent estimations of the atmospheric states. Since these form a non-linear system of partial differential equations, it is not possible yet to derive an analytical solution, and an approximate solution must be developed (Bjerknes, 1904). The solution is a function of the initial conditions which need to be provided in advance. The first attempt to numerically integrate the equations was performed by hand and it resulted in an unrealistic estimation (Richardson, 1922). It was necessary considerably progress in our understanding of the atmospheric dynamics and numerical methods before a successful estimation was achieved (Charney et al., 1950).

The numerical algorithms used to simulate the global behavior of the atmosphere are known as General Circulation Models (GCMs). The GCMs have been shown to reproduce the latitudinal eastward-westward-eastward surface wind alternation, and have contributed to make progress in our understanding of the general circulation of the atmosphere (Phillips, 1956; Lewis, 1998). The integration of the hydrodynamic equations over the whole atmosphere requires large computational resources which severely limit the spatial discretization. Even with the large computer development experienced during the last decades, the typical GCMs horizontal resolution is on the order of 1^o lat x lon (111 km at the equator). Hence, GCMs can potentially provide adequate results at homogeneous regions but their results must be considered carefully as the surface heterogeneity increases. This is specially the case of complex terrain regions, wherein the surface wind field can present a high variability as a consequence of the strong modifications that the orography produces over the large scale flows.

Theoretically, this limitation of the GCMs could be avoided by increasing the horizontal resolution to capture regional details. However, the computational cost necessary to approximate a solution would be prohibitive. It is therefore necessary to look for additional procedures to obtain regional information. With this aim, downscaling techniques based on the good performance of the GCMs to reproduce large scale fields were developed (Giorgi and Means, 1991; von Storch, 1995; Hewitson and Crane, 1996; Wilby and Wigley, 1997). The downscaling approaches can be divided into statistical and dynamical depending on the procedures used to derive the regional information.

The statistical approaches establish empirical relationships between large scale fields, usually outputs of GCMs, and observed variables during a calibration period (Zorita et al., 1992; von Storch et al., 1993; Wilby et al., 1998). The relationship is then used to estimate the regional variability in the absence of observations. Quite different approaches based on either linear or non-linear techniques can be used to establish the statistical relationship between the large scale and the regional scale. All the statistical

downscaling approaches suffer from the assumption that the relationship encountered during the calibration period will hold on for other temporal periods. In addition, these techniques are able to provide estimations only at those places where observations are available. On the contrary, the statistical downscaling presents the advantage of requiring few computational resources to derive the regional variability once the relationship has been found. The performance of statistical downscaling approaches has been extensively explored for precipitation and temperature variables (e.g. Kim et al., 1984; Wigley et al., 1990; Noguer, 1994; González-Rouco et al., 2000; Huth, 2002; Xoplaki et al., 2003a,b, 2004; Vrac et al., 2007; Paul et al., 2008; Busuioc et al., 2008; Li and Smith, 2009). On the contrary, an exhaustive analysis of their performance with wind related variables was seldom attempted. A pioneering exercise used linear methodologies based on canonical correlation analysis to reconstruct a wind climatology for the last century at several locations in the North Atlantic area (Kaas et al., 1996). Another study used statistical downscaling to reconstruct the wind climatology of several buoys in western Canada (Faucher et al., 1999). Other statistical downscaling exercises have dealt only partially with wind variables (e.g. Gutiérrez et al., 2004; Cheng et al., 2007) or use them to estimate potential wind speed variations under climate change scenarios (Bogardi and Matyasovszky, 1996; Sailor et al., 2000; Pryor et al., 2005b,c, 2006; Najac et al., 2009).

The dynamical downscaling techniques derive regional variability using physical principles similar to those used by GCMs (Krishnamurti and Bounoua, 1995; Pielke, 2002). However, the atmospheric evolution is simulated only over a limited area of the Earth, which allows to reach horizontal resolutions of a few kilometers. Increasing the spatial resolution provides a more accurate representation of the surface heterogeneity and a better representation of physical processes (Stensrud, 2007), which should be reflected in a more accurate simulation (Mass et al., 2002). The limited area models or mesoscale models make use of GCM outputs to obtain the initial and boundary conditions necessary to approximate a solution. A number of different models are available to perform the downscaling (e.g. Black, 1994; Grell et al., 1994; Bubnova et al., 1995; Cotton et al., 2003; Skamarock et al., 2005).

The limited area models were first used for numerical weather prediction or to analyze mesoscale phenomena (Physick, 1988; Warner, 1989). For those purposes, only one or two days of simulation were required. The idea of using limited area models to understand the regional impacts of the global climate system was proposed by Dickinson et al. (1989). Shortly after that study, longer regional climate simulations spanning a few months (e.g. Giorgi and Bates, 1989; Giorgi, 1990; Marinucci et al., 1995; Walsh and McGregor, 1995) and multi-year simulations were accomplished (e.g. Giorgi et al., 1993; Christensen et al., 1998). The field experimented an impressive development in successive years (McGregor, 1997; Wang et al., 2004; Giorgi, 2006). Several studies have evaluated the performance of multi-year simulations to reproduce the surface wind field (Buckley, 2004; Conil and Hall, 2006; Walter et al., 2006; Kanamitsu and Kanamaru, 2007). However, long simulations at a very high spatial resolution (a few km), necessary to provide a relatively accurate representation of the details of complex orography, are still computationally demanding and thus have not been so far extensively used. Usually, a short temporal period of a few months or specifically the duration of an observational campaign are simulated in order to dimension the computational requirements (e.g. Georgelin and Richard, 1996;

Masson and Bougeault, 1996; Zhong and Fast, 2003; Cairns and Corey, 2003; Rife et al., 2004; Rife and Davis, 2005; Zagar et al., 2006; Soriano et al., 2006; Pérez-Landa et al., 2007). These studies contribute to understand the influences that orography exerts over the surface circulations or to evaluate the performance of the simulation under specific synoptic situations. However, it is also necessary to evaluate the capability of dynamical downscaling approaches to reproduce the long term characteristics of the surface flows over complex terrain regions.

Considerable progress in our understanding of the behavior of both statistical and dynamical downscaling approaches to reproduce surface wind circulations seems to be necessary before using these techniques in routine activities with a certain degree of confidence.

1.2 Objectives and rationale

This study aims to analyze the capability of a multi-year dynamical downscaling accomplished at a high spatial resolution to reproduce the surface wind variability and climatology over a complex terrain region. The analysis is performed at daily time scales. This decision filters out the effects associated with thermally driven circulations, and allows to focus on the influence of dynamically driven circulations associated with the atmospheric dynamics and its interaction with the topography of the region. In order to provide an accurate evaluation, the surface wind variability and climatology is first analyzed with observations, and the knowledge gained is subsequently used to evaluate the behavior of the numerical simulation. With this aim, the observational analysis are oriented to understand the surface circulations over the area and at a time provide an appropriate framework to evaluate the performance of the dynamical downscaling. Hence, there is a twofold objective in the present study: i) analyze the surface wind variability and climatology over a complex terrain region; ii) to evaluate the performance of a dynamical downscaling approach to reproduce the observed behavior.

The downscaling is accomplished at a very high spatial resolution over the area of study in order to provide a relatively accurate representation of the topographic features. In addition, the simulation spans a relatively long temporal period to provide a robust evaluation of the large scale influence on the downscaled surface circulations. We are not aware of previous studies analyzing the dynamical downscaling performance to reproduce the surface wind behavior at such a high spatial resolution and over such an extended temporal period as the one described herein.

There is not an obvious reason to justify the selection of dynamical downscaling approaches in detriment of the statistical ones, since both methodologies present their drawbacks and advantages and both deserve being evaluated. It can be argued that for the analysis of any meteorological field, the dynamical approach presents certain advantages over the statistical methodologies due to the higher spatial coverage provided, physical consistency and relative independence of observations in the building of the downscaling strategy.

The comparison of numerical simulations with observations is usually referred to as “verification” or “validation”. However, verification or validation of numerical models of

natural systems can be argued to be impossible (Oreskes et al., 1994; Oreskes, 1998). The simulations can be compared to observations to confirm that the simulation is appropriated for a particular case study, but nothing ensures that it will continue to work in other situations. The higher the number of comparisons when the simulation shows an adequate behavior, the more likely that the simulation would perform appropriately in other situations. In this sense, the models can only be evaluated in relative terms (Oreskes et al., 1994). In addition, the evaluation of a given simulation meets some uncertainties when comparison with local information is involved (von Storch, 1995). Usually, *in situ* observations are compared against the nearest simulated grid points (Cox et al., 1998; Hanna and Yang, 2001; Buckley, 2004). However, two main reasons are worth to stress as responsible for the uncertainty introduced in this particular type of comparison. First, the simulated variables represent averaged quantities over volumes with homogeneous properties due to the spatial discretization and the Reynolds averaging. Therefore, its comparison to *in situ* observations is controversial, mostly at those locations that are considerably affected by local features. Second, the discretization smooths the complexity of orography and surface physical properties. This can potentially lead to a situation in which the simulated volume, that includes the actual location with observations, may not be the most suitable one to represent the observational variability; instead, nearby volumes may be more appropriate. This last problem will be referred to hereafter as the *representativeness error*. To mitigate such adverse effects, Reid and Turner (2001) compared the volume averaged quantities simulated at a coarse horizontal resolution (40 km) against averaged observations within grid cells representing subjectively defined subregions. The averaging of the observational series to obtain regional ones filters out local variability (noise), thus enhancing the regional signal. This led to stronger relationships between model and data than in the traditional nearest grid point comparison.

Reid and Turner (2001) highlight the advantages of an evaluation performed at the regional scale but their definition of subregions is however subjective. As a step forward towards a more objective evaluation, this study uses an automated approach to identify the regions with similar wind variability without any *a priori* knowledge of the wind behavior over the area of study. Indeed, the regionalization allows to analyze the spatio-temporal wind variability over the region (Chapter 4). The identification of areas of internally coherent wind behavior also provides an appropriate framework to evaluate the performance of the dynamical downscaling in reproducing the surface wind variability (Chapter 5). In particular, averaging the time series to obtain the regional ones damps local effects in the observations as well as random *representativeness errors* in the simulation leading to more appropriate time series for comparison. Additionally, the higher spatiotemporal coverage provided by the downscaling is used to analyze the wind variability where and when observations are not available.

The analysis of the climatological surface circulations over the area provides complementary information to the previous wind variability analysis. The climatological surface flows are herein studied through the identification of wind patterns (WPs) that characterize local circulations (Chapter 6). For this purpose, the observational wind fields with similar structures are classified together. The classification allows to identify the prevailing surface circulations and their relation to regional climate. The sea level pressure (SLP) fields are also subjected to a classification procedure in order to identify the

typical synoptic circulations or pressure patterns (PPs) that affect the area of study. The relationship between the PPs and the already identified WPs allows to understand the synoptic mechanisms responsible for regional atmospheric variability. The classification of the wind fields into WPs is also used to evaluate the performance of the dynamical downscaling in reproducing the climatological surface flows over the region (Chapter 7). In addition, the knowledge of the synoptic situations that can produce a given WP, their associated PPs, allows to discriminate the downscaling performance in reproducing the climatological surface flows under their potential synoptic forcings. Again, the evaluation is performed at the subregions with homogeneous wind behavior previously identified due to the advantages that a regional evaluation provides in comparison with the evaluation at the sites (see above). Also, the length of the simulation allows for a large number of days within each WP class. This feature contributes to increase our confidence to generalize the results obtained during the evaluation (recall the previous discussion of Oreskes, 1994, and Oreskes, 1998).

The Comunidad Foral de Navarra (CFN), a relatively complex terrain region located in the Northeast of the Iberian Peninsula (IP, Fig. 1.2) was selected for this case study. The CFN is characterized by strong wind conditions which have favored the development of wind energy facilities in recent years (Fairless, 2007; García-Bustamante et al., 2008, 2009). The CFN presents a reasonably dense surface meteorological network including wind observations available since 1992 that has made it possible to undertake various analysis of surface wind over the area (García et al., 1998; Torres et al., 1999, 2005; Jiménez et al., 2008, 2009b). Therefore, this region constitutes, given the existence of complex orography and good availability of data, an interesting case to evaluate the performance of the numerical simulations.

The Weather Research and Forecast (WRF) modeling system is the mesoscale model used to perform the dynamical downscaling (Skamarock et al., 2005). The whole available observational period from 1992 to 2005 is simulated in this work at a spatial resolution of 2 km (Chapter 3). This provides a long simulation with both a high spatial and temporal resolution, comparable to that of the observational dataset. The value of the WRF simulation should be viewed in the context of the computational costs of running mesoscale models. We are not aware of previous studies that evaluate the performance of mesoscale simulations of these characteristics.

1.2.1 The selected region

The IP surface circulation is controlled by the semi-permanent subtropical high pressure center over the Azores Islands (Zimmerschied, 1949; Sahsamanoglou, 1990). In winter, it is usually centered at lower latitudes (Davis et al., 1997) and thus the whole IP is often affected by zonal circulations from the West combined with the perturbations originated by the Polar Front. In summer, the subtropical high pressure center expands towards higher latitudes, blocking the western circulations over the peninsula except for the northern region (Davis et al., 1997). In addition, the relative displacements and changes in intensity of this semi-permanent subtropical high allow that a variety of air masses with distinct physical properties influence the circulation over the IP at regional and local scales. Polar continental air masses from the semi-permanent Siberian High, tropical continental

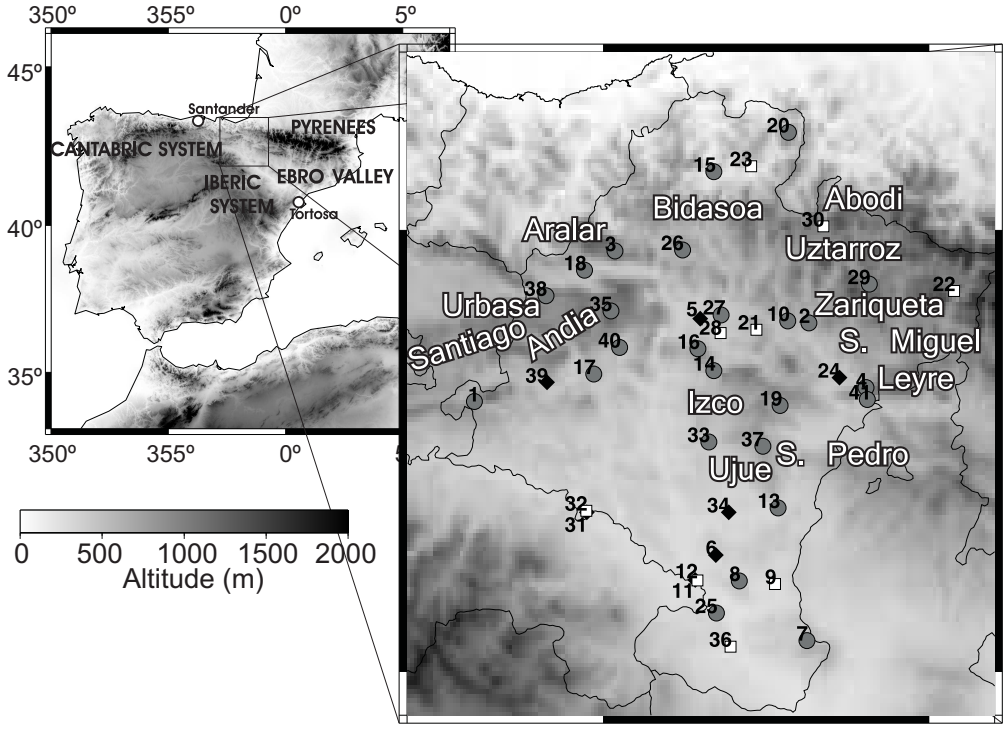

Fig. 1.2: Location of the CFN and meteorological stations (symbols). Left panel highlights the most relevant geographical features of the IP surrounding the CFN and the right panel illustrates some regional details within the CFN. Thin lines in the right panel represent political boundaries. Circles, diamonds and squares represent the location of the GN, INM and RN observational stations, respectively. See Table 2.1 for an specific station description.

air masses from the semi-permanent Sahara Low, arctic and polar maritime air masses coming through the North and Northwest of the Peninsula and subtropical or tropical maritime air masses coming from the West or Southwest are some of the well known synoptic flows which affect the IP (Font, 2000). The different properties of the advected air can lead to dust intrusion by southern flows coming from the Sahara desert (Rodríguez et al., 2001), advection of moisture by the western and southeastern flows (Trigo and DaCamara, 2000) or favour the occurrence of minimum extreme temperature episodes over the peninsula (Prieto et al., 2004). In addition, if air masses remain long enough over the Mediterranean sea, they can potentially enlarge their water content and they could move towards the IP penetrating from the East or Southeast. This Mediterranean air masses have been shown to prevail during summer where they play an important role in air pollution dynamics (Millán et al., 2003). The interaction of the large variety of air masses with trajectories intercepting the IP, and a particular complex topography (Fig. 1.2) produces a range of differentiated regional climates.

The rough orography of the CFN region produces an interesting degree of interaction with large scale dynamics. The region is broadly limited by two large mountain systems: the Iberian System in the South and the Pyrenees in the North, which merge westwards with the foothills of the Cantabrian Mountains. Between them, the Ebro Valley crosses the region from Northwest to Southeast towards the Mediterranean. A closer look at the CFN (see zoomed area in Fig. 1.2) reveals a rather complex array of smaller mountain systems and valleys. The North is dominated by the Pyrenees with the large Bidasoa mountain lines or the smaller ridges of Abodi, Uztarroz, San Miguel, Zariquieta and Leyre. To simplify, the latter will be referred to as the *northern mountains*. The western and northwestern boundaries are outflanked by the Aralar mountain lines as well as the Urbasa, Santiago and Andía which will be herein labelled as the *western mountains*, unless treated specifically. The center and eastern side of the CFN is punctuated with the mountain systems of Izco, San Pedro and the Ujué peak which will be referred to as the *eastern mountains*. Finally, the South of the CFN is dominated by the lower lands of the Ebro Valley. Virtually parallel to the Ebro valley, smaller valleys line up in a NW-SE direction south of the *northern mountains*. Similarly, several of the mountain systems and valleys West and East of the *eastern mountains* favor the wind channeling in the NW-SE and in the N-S direction.

The CFN regional climate presents an important contrast between the North and the South since the orography and the specific interactions with the different air masses induce rather unlike physical conditions to the flow. The northern part is exposed to mild and moist air from the Atlantic ocean during the whole year (García and Reija, 1994), while the southern part is dominated by the presence of the Ebro valley, whose bounding ridges isolate it from the Atlantic circulations and give rise to a dry region where winds are channeled and intensified (Biel and García, 1962). Northwestern cold dry winds blowing down-valley are known as *Cierzo* and southeastern warm moist winds blowing up-valley as *Bochorno* (García, 1985).

1.3 Structure

The study is organized as follows. Chapter 2 describes the observational data and the procedures applied to them in order to suppress and correct doubtful records. Contrary to the relatively well referenced quality control (QC) of temperature and precipitation variables, literature for wind variables is sparse (DeGaetano, 1997, 1998; Graybeal, 2006). In particular, there is not any previous work devoted to ensure the quality of high frequency wind observations from automated weather stations as those presented herein. Chapter 3 briefly describes the WRF model and the configuration adopted to perform the dynamical downscaling. The subsequent chapters analyze the surface wind over the region through observations and in the numerical simulation.

The daily mean wind variability over the CFN is analyzed in Chapter 4. This is performed by grouping together those observational sites with similar temporal wind variability. The regionalization allows for the identification of areas of coherent wind behavior. Results from the classification are used in Chapter 5 to evaluate the performance of the dynamical downscaling in reproducing the surface wind variability.

The climatological surface flows over the CFN are analyzed in Chapter 6. Both, the classification of SLP and surface wind data into classes of the most common synoptic pressure patterns (PPs) and surface wind circulation patterns (WPs) is provided as well as the associations between them. The performance of the dynamical downscaling to reproduce the climatological surface circulations, the WPs, is presented in Chapter 7. The evaluation focuses on the influence of the orography and the large scale effects on the downscaled surface flows.

Finally, Chapter 8 presents the conclusions and an extended discussion of the work.

2

Observational dataset: quality control and correction of biases*

I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be.

Lord Kelvin, 1883.

The surface wind variability and climatology over the CFN is herein analyzed using observations acquired at an automatic weather station network. This chapter presents the observational dataset and the procedures applied to the original records in order to improve their quality.

2.1 Background and motivation for quality control and correction of biases

Meteorological observations are essential to understand the physical processes that affect the regional and global climate. The quality of the data used in a given study affects, in many ways, the reliability of the conclusions reached. It is therefore very important to keep observational datasets as error-free as possible. According to Gandin (1988), the errors associated with meteorological records can be classified in three main groups: random, systematic and rough errors. Random errors are intrinsic to the measurement definition, an approximation of the real atmospheric state, and therefore these errors are unavoidable. Systematic errors are associated with a more or less persistent factor that introduces a certain bias in the registered values. Biases can be produced by a recalibration of the recording instrument, changes in exposure or type of the sensor, changes in the observing time, etc (e.g. Karl et al., 1986; Parker et al., 2000; Pielke Sr. et al., 2007b). The techniques developed to identify and correct for these biases in time series are known as homogeneity adjustments (Alexanderson, 1986; Karl and Williams,

* The main contents of this chapter are included in:
paper prepared for sumbision (reference to be updated in the last version of this manuscript)

1987; Peterson et al., 1998a). Rough errors are associated with a misfuncton of the sensor or mistakes introduced during the data processing, transmission, reception or storage. Recognition and suppression of this third type of erroneous records are the objectives of any QC process (Gandin, 1988).

The identification and correction of systematic errors (or biases) and the suppression of rough errors from the climatic datasets is a tedious and time-consuming task. To optimize the resources required in this process, it is particularly important to understand the routine observational activities as well as the sporadic incidents that have affected the acquisition of measurements. This information is known as metadata (WMO, 2008) and it facilitates the identification of suspicious observations through a specific design of the procedures applied to assess the veracity of the records.

The surface meteorological observations can be acquired either manually or automatically from the weather stations. The first atmospheric records available for climatic research came from manually controlled stations. Temperature and precipitation are by far the most studied variables and therefore the systematic errors associated with their measurement are those best understood (e.g. Karl et al., 1989). In spite of our understanding of the potential biases, its detection is always complicated and even nowadays new biases appear on the climatic datasets (Pielke Sr. et al., 2007a; Thompson et al., 2008).

The main focus of the QCs was the outlier detection (Filippov, 1968; Grant and Leavenworth, 1972). Outliers could be associated with an anomalous temporal or spatial variability and specific checks to detect each one of them have been developed in past QC exercises (Eischeid et al., 1995; Peterson et al., 1998b; González-Rouco et al., 2001). In the last years, temperature and precipitation records are being analyzed on a daily or even hourly basis and the complexity of the strategies used in the QCs and bias corrections was considerably increased (Reek et al., 1992; Kunkel et al., 1998; Vincent et al., 2002; Wijngaard et al., 2003; Kunkel et al., 2005; Brunet et al., 2006). Current research has mainly focused on comparing the performance of different approaches to identify rough errors (Hubbard et al., 2005, 2007; You et al., 2007; Durre et al., 2008). The knowledge gained through the correction of errors associated with temperature and precipitation variables has recently favored the correction of errors associated with other meteorological variables which were manually acquired like surface wind, pressure or humidity (e.g. Graybeal et al., 2004a,b; Feng et al., 2004; Wan et al., 2007).

The introduction of automatic weather stations has allowed to increase the temporal resolution of sampling. In addition, observations are taken at regular time intervals along the day. This has eliminated some biases of the manual observations associated with changes in the timing of observation (Wu et al., 2005). The new automatic protocols have logically introduced a new variety of systematic and rough errors that are still poorly understood. Perhaps the best choice to deal with these unknown errors is to mitigate their potential influence by performing routine maintenance operations on the weather stations as well as applying daily QCs to the registered records (Shafer et al., 2000; Fiebrich et al., 2006; McPherson et al., 2007). However, not all the automatic networks follow these desirable practices and sometimes the observations are just stored without applying any QC (Tucker, 1997).

In recent years there is an increasing demand for high quality wind observations in a wide range of applications: the evaluation of wind energy resources, the transport and dispersion of pollutants within the frame of air quality studies, the analysis of extreme wind events by insurance companies or for the design of buildings and structures, etc. Despite this necessity for data of quality, QCs that specifically deal with wind related variables from either manual or automatic origin are scarce (Wade, 1987; DeGaetano, 1997, 1998; Graybeal, 2006). Our understanding of the potential errors that are latent on the climatic wind datasets is therefore very limited and hampers the creation of the high quality datasets demanded.

In an early work, Wade (1987) analyzed wind data taken at a dense network of automatic weather stations during a three month campaign. He identified biases introduced on the wind speed records due to obstacles nearby the station as well as biases in the wind direction as a consequence of a wrong fixation of the wind vane offset. More recently, DeGaetano (1997) developed a complex QC for hourly records manually obtained during almost 40 years at 41 stations in the eastern USA. A small percentage of the records were considered suspicious indicating a good quality of the data, but he did not rule out the possibility that other potential errors could still be present on the dataset. DeGaetano (1998) went further on by analyzing biases in the dataset. He found a tendency to report even records, rounding problems and a distinct treatment of the calm depending on the institution that maintained the weather station. Graybeal (2006) analyzed the relationship of extreme wind variables with the daily mean wind speed to establish a relationship that would be helpful to evaluate the quality of the extreme wind records. Aside from these studies, errors associated with wind records have only been analyzed as part of more ambitious programs that involved the correction of errors in other meteorological variables, and therefore without going into depth in the specific problems related to these type of measurements (Meek and Hatfield, 1994; Shafer et al., 2000; Graybeal et al., 2004a; Feng et al., 2004).

This chapter summarizes the QC and the correction of biases applied to the wind observations taken at the 41 automatic weather stations over the CFN. The observations available were collected over a thirteen year period at a high temporal resolution (10-min or 30-min). The QCs applied to the data during the acquisition period were very limited or inexistent and thus, the potential number of erroneous records could be high. In this sense, this work does not only try to contribute with the creation of a wind dataset of improved quality, but also reports about potential errors that could be present in other wind datasets.

2.2 Observational wind data

The wind speed and direction measurements were collected from 1^{st} January 1992 to 7^{th} October 2005 at 41 automatic weather stations distributed over the CFN (Fig. 1.2, Table 2.1). The stations are managed by three different institutions: Gobierno de Navarra (GN, circles in Fig. 1.2), Agencia Estatal de Meteorología (AEMET, though still known as Instituto Nacional de Meteorología, INM, at the time of the data adquisition; squares in Fig. 1.2) and Riegos de Navarra (RN, diamonds in Fig. 1.2). The observations of the

GN stations are acquired either via modem, with three calls every day, or in real time via radio-trunking. This institution, GN, receives the observations from AEMET and RN by a daily ftp and is responsible to store and maintain the whole dataset. Information concerning maintenance operations of the weather stations is quite limited. Only the operations performed to the wind sensors of the GN stations from 2000 to the end of 2002 were provided.

All stations recorded the wind speed in ms^{-1} and the wind direction in degrees respect to North. However, two different wind sensors were used to acquire observations. The wind sensor of GN and RN is able to measure wind speeds up to 60 ms^{-1} and presents a cut-in wind speed of 1.0 ms^{-1}. The anemometer employed by AEMET is able to register a maximum wind speed of 50 ms^{-1} and presents a cut-in wind speed of 0.3 ms^{-1}. The temporal resolution of the records is 10 minutes except for the stations managed by RN that after March 2004 provided averages of the wind observations every 30 minutes.

Another source of discrepancy between the various datasets is the treatment of calms. Some datasets assign the zero wind direction value to zero wind speed records, keeping 360° for northern winds. Other datasets treat independently the wind speed and direction measurements without a special attention to calm records. The original wind dataset provided by GN contained both treatments of calm records. This was due to the different criteria selected by each institution as well as to changes in the criteria adopted through time. To avoid this bias in the dataset, it was decided to adopt the same calm criteria for all the records. Since there are periods with zero wind direction values for calm winds and its transformation to regular wind direction values is not obvious, the criteria that assigns zero wind direction values to zero wind speed records was finally adopted.

The following sections describe the procedures used to improve the quality of the dataset provided. The first subsection explains the procedures applied to suppress rough errors, the QC (Section 2.2), and the second one deals with the systematic errors, the biases (Section 2.3). Examples of the kind of problems that are addressed are also provided. Section 2.5 and 2.6 report on the corrections applied and Section 2.7 describes the impacts that the corrections produced on the main statistics of the time series.

2.3 Methodologies: quality control

The QC applied herein consists of a sequential application of several checks attempting to identify questionable data in both wind speed and wind direction time series. Tests are arranged in a logical sequence to effectively assess the reliability of observations attending to the kind of problem. There were some checks in the QC that directly flagged observations as invalid while there are others that reported questionable observations for a later inspection introducing a certain degree of subjectivity in the QC. This is an undesirable characteristic that was mitigated by designing the checks in such a way that minimized the number of periods to be manually inspected. No attempt was made to fill up the missing values introduced on the time series.

The flux diagram of the QC is shown in Figure 2.1. The wind speed and wind direction data are assessed independently although the majority of the steps; the ones in the central column of the QC diagram, are common for both variables. The tests are subdivided

Table 2.1: Code of the meteorological station as in Figure 1.2, name, longitude, latitude, altitude, sensor height above ground level and installation date (dd/mm/yy).

Num.	Name	Long. (^{o})	Lat. (^{o})	Alt. (m)	Height (m)	Inst. date
1	Aguilar de Codés	-2.394	42.614	731	10	01/03/92
2	Aoiz-Agoitz	-1.369	42.792	530	10	01/03/92
3	Aralar	-1.963	42.954	1393	10	01/03/92
4	Arangoiti	-1.194	42.646	1353	10	01/01/92
5	Arazuri	-1.702	42.801	396	2	04/02/00
6	Bardenas-barranco salado	-1.654	42.265	300	2	01/03/98
7	Bardenas-loma negra	-1.375	42.071	646	10	01/03/92
8	Bardenas-Nstra Sra. del Yugo	-1.582	42.206	472	10	01/01/92
9	Bardenas-polígono de tiro	-1.473	42.200	295	10	01/05/97
10	Beortegi	-1.434	42.796	580	10	01/05/97
11	Cadreita-Riegos	-1.717	42.209	268	2	01/03/98
12	Cadreita-INM	-1.710	42.208	268	10	01/03/92
13	Carcastillo	-1.463	42.372	340	10	01/03/92
14	Carrascal	-1.660	42.683	560	10	01/01/92
15	Doneztebe	-1.660	43.132	138	10	01/06/99
16	Perdón	-1.709	42.733	1024	10	01/03/92
17	Estella-Lizarra	-2.028	42.676	480	10	01/03/92
18	Etxarri-Aranatz	-2.057	42.910	507	10	01/03/92
19	Getadar	-1.457	42.605	710	10	01/05/00
20	Gorramendi	-1.432	43.220	1071	10	01/05/92
21	Ilundain	-1.529	42.777	542	10	01/03/92
22	Isaba	-0.923	42.864	843	10	01/07/92
23	Lekaroz	-1.545	43.144	182	10	01/03/92
24	Lumbier-Ilumberri	-1.275	42.668	484	2	05/05/00
25	Montes del Cierzo	-1.652	42.133	310	10	01/07/98
26	Oskotz	-1.756	42.956	562	10	01/03/99
27	Pamplona-Larrabide	-1.638	42.810	450	10	01/01/97
28	Pamplona-Noain	-1.639	42.769	461	10	01/04/92
29	Remendia-Salazar	-1.184	42.879	1047	10	01/10/01
30	Roncesvalles-Orreaga	-1.325	43.009	940	10	01/03/92
31	Sartaguda-Riegos	-2.050	42.363	310	2	01/03/98
32	Sartaguda-INM	-2.051	42.366	310	10	01/03/92
33	Tafalla	-1.676	42.522	415	10	01/03/92
34	Traibuenas	-1.614	42.363	312	2	14/04/99
35	Trinidad de Iturgoien	-1.975	42.819	1222	10	01/01/92
36	Tudela	-1.608	42.057	295	10	01/03/92
37	Ujué	-1.510	42.513	826	10	01/01/92
38	Urbasa	-2.175	42.853	890	10	01/10/01
39	Valdega	-2.172	42.657	469	2	01/05/01
40	Villanueva del Yerri	-1.949	42.736	498	10	01/01/98
41	Yesa	-1.190	42.618	489	10	01/03/92

into three main groups: 1) checks for the detection of manipulation errors associated with the storage and maintenance of a dataset; 2) limits consistency checks to remove records outside of the allowable range of variation; 3) and checks to ensure the temporal consistency of the individual time series by assessing records with abnormally low/high variations. The following sections describe the various cases.

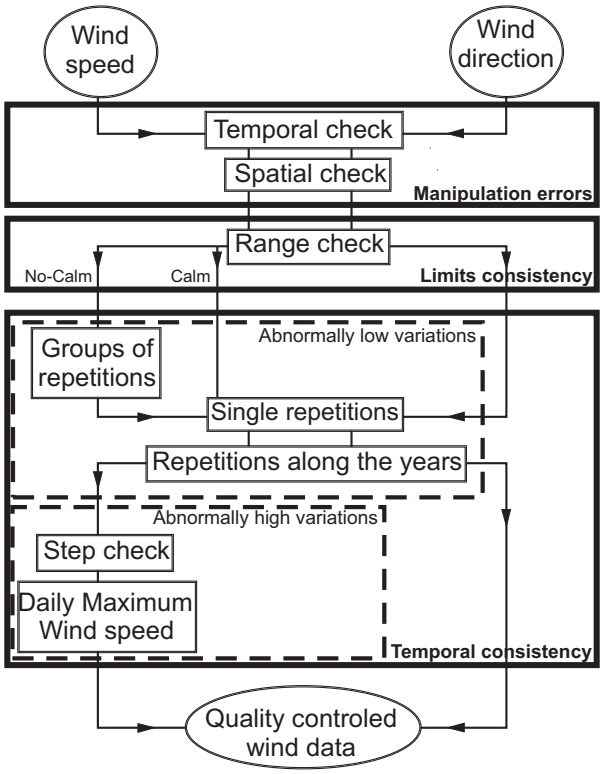

Fig. 2.1: Diagram of the QC. The name of each check applied to the data is represented within a thin rectangle. The three rectangles in bold comprise the checks associated with manipulation errors, limits consistency and temporal consistency. The last one contains two additional rectangles (dashed) that comprise the checks that deal with abnormally low and high variations in the records.

2.3.1 Manipulation errors

Firstly, wind speed and direction observations were screened to detect manipulation errors. These type of errors are associated with the management and storage of the data. The checks applied herein look for a temporal or spatial misplacement of observations

that can coincide with other measurements in a given time series or with measurements at other stations.

The temporal assessment attempts to detect periods with repeated observations within the measurements at a single station. It requires the choice of a time step to discretize the time series in order to define the periods to check. The length of the temporal period used to discretize the time series should be that one at which the data is received and stored. In this particular wind dataset, the frequency in the data acquisition was eight and twenty four hours (see section 2.2) and therefore these were the time steps selected. Thus, each period was compared against the rest looking for coincident measurements, and verifying their variances were higher than zero. The date of the period and the code of the station were reported for a posterior inspection. One of these cases can be observed in Figure 2.2a. The wind speed observations from station Lekaroz (23) show several days with a strong diurnal cycle due to its natural variability but, there are identical measurements in the first and the second day of May. These days also present coincident measurements in the wind direction time series (not shown), suggesting a wrong temporal assignment of observations from one day to the other. Since this type of coincidence is very unlikely to happen, both wind speed and wind direction measurements of the two days were eliminated and flagged as missing values.

The spatial check for manipulation errors allows to detect coincident measurements between stations. It reports the date of the coincidences and the code of the stations if they present the same measurements for more than eight hours, which is the shortest period in the data acquisition, and if their variance is higher than zero. The imposition of variance higher than zero, both in the temporal and in the spatial assessments, is to avoid periods with constant measurements. These will be treated in posterior checks. A typical coincident interval in the measurements of two stations is displayed in Figure 2.2b. The wind speed time series from station Aguilar de Codés (1) and station Aoitz-Agoitz (2) are coincident during one day. Wind direction presents a similar behavior with coincidences in the same day as wind speed (not shown). The identical wind observations were finally removed from the time series of both stations.

2.3.2 Limits consistency

After correcting for manipulation errors, a range check to remove unrealistic observations is applied (Fig. 2.1). The range checks impose high and low limits for the possible values of one variable and consider invalid the observations outside the range allowed. The threshold values could be based either on static or dynamic climatic extremes or on the response ranges of the given sensors (Meek and Hatfield, 1994). In this QC, the limits are fixed attending to the anemometer characteristics. Hence, the low wind speed limit is fixed to 0.0 ms^{-1}, and the high limit to 60.0 ms^{-1} for the RN and GN stations and to 50.0 ms^{-1} for the AEMET stations. For the case of wind direction, a low limit of 0° and a high one of 360° are selected. A period with wind direction records that exceed these limits of variability registered at station Arangoiti (4) is shown in Figure 2.3. The measurements reveal northwestern flows, close to 360°, and suddenly they exceed this limit and show values around 450° and 540°. After that, the records show northwestern flows again. The causes that originate these erroneous records are not obvious. The wind

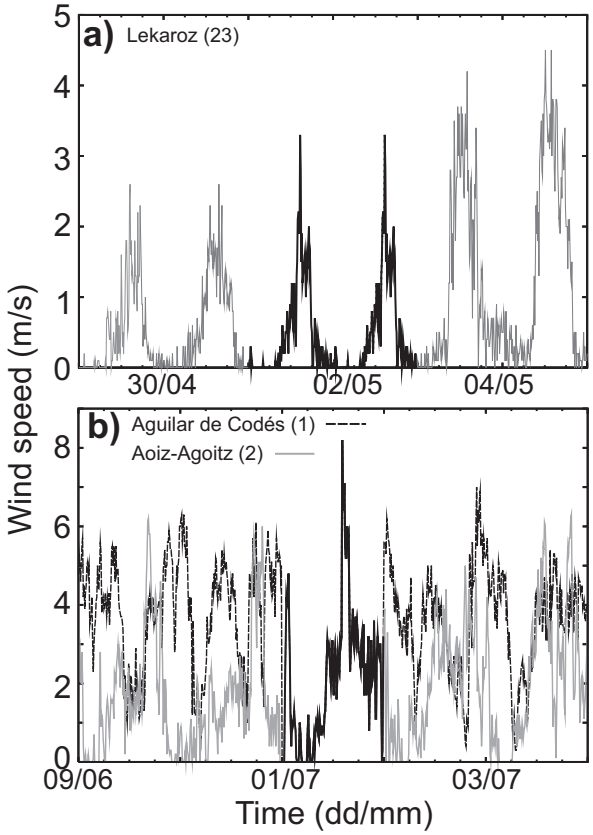

Fig. 2.2: Examples of wind speed records in two cases of identical observations: a) records from a period registered at Lekaroz (23) associated with a potential temporal misplacement of observations, and b) records from Aguilar de Codés (1) and Aoiz-Agoitz (2) during a potential spatial misplacement of observations. The black solid lines highlight the coincident records.

speed records associated with this temporal period were analyzed looking for potential explanations, but nothing suspicious was found. The records higher than 360^o were finally removed.

2.3.3 Temporal consistency

The next step of the QC is to ensure the temporal consistency of the time series (Fig. 2.1). Wind speed checks are divided into two groups, the first one looks for periods with abnormally low variations and the second one for abnormally high variations. For the case of wind direction time series, only temporal consistency checks that look for periods with abnormally low variations are applied. This is because the effects of turbulence are of

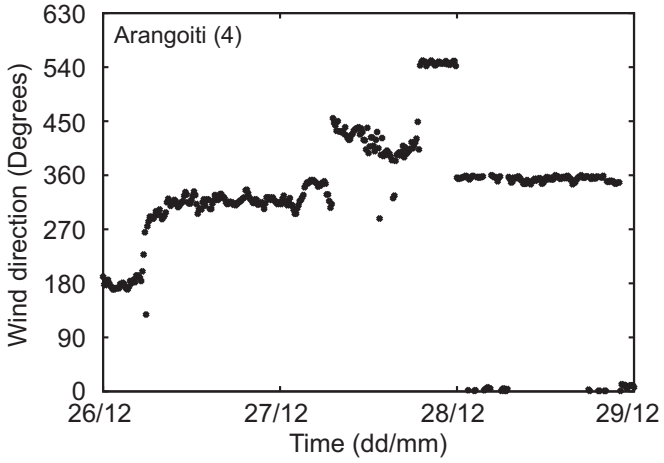

Fig. 2.3: Wind direction records registered at Arangoiti (4).

relevance in the timescales considered, 10-min and 30-min, and can produce relatively fast and large changes within which even a reversal in wind direction between two consecutive records could be argued to be valid.

2.3.3.1 Abnormally low variations

Abnormally low variations are those periods which present remarkably long constant observations. In the case of wind speed, a difference is made between constant periods of wind speeds equal or higher than 1.0 ms^{-1} and those of wind speeds lower than 1.0 ms^{-1}. This allows for a separate treatment of repetitions associated with calm periods and repetitions related to other phenomena. This differentiation was performed in previous wind QC (DeGaetano, 1997). The wind speed limit that separates between calm and no-calm should be the cut-in wind speed of the anemometer, since lower wind speeds measurements result from averaging over one or more zero records. The selection of 1.0 ms^{-1} is in agreement with the cut-in wind speed of the GN and RN sensors (see section 2.2). The possibility of selecting a threshold of 0.3 ms^{-1} for AEMET stations was considered since this is the cut-in wind speed of its sensors. However, to maintain this theoretical cut-in wind speed of 0.3 ms^{-1} it would have been necessary that strict maintenance operations had been performed over the sensors in order to mitigate its deterioration by the effects of adverse meteorological conditions (Fiebrich et al., 2006). Since there are not metadata available to evaluate if maintenance operations were adequate, and some preliminary inspections revealed that the degradation of the sensor accuracy could indeed occur, the limit between calm and no-calm wind speeds for these stations was fixed, as in the rest of stations, to 1.0 ms^{-1}.

The distinction between wind speeds associated with calm or no-calm periods to deal with abnormally low wind speed variations causes a separation in the flow of the QC (Fig.

2.1). Repetitions of wind speeds equal or higher than 1.0 ms^{-1} appear, most of the times, in clusters. These periods of dubious quality are individually analyzed and suppressed if considered appropriate. A typical case is displayed in Figure 2.4. Repetitions of wind speed values over twenty days can be observed oscillating around 21 ms^{-1}. Furthermore, there exist periods with missing values before, between and after the repetitions. The veracity of this group of measurements is highly questionable and thus the whole period was considered invalid.

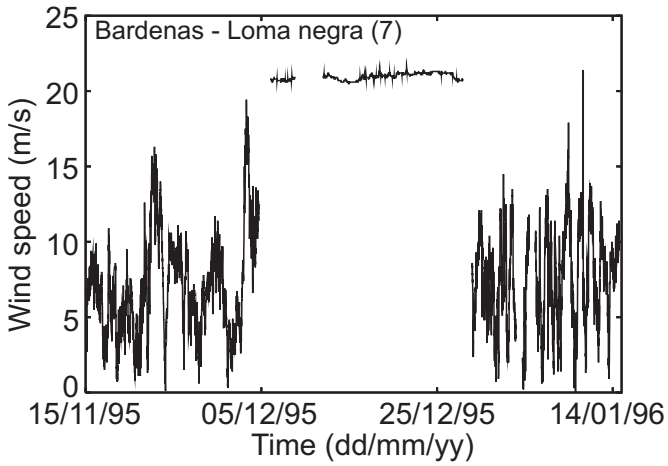

Fig. 2.4: Observations corresponding to a period of the wind speed measurements from Bardenas-Loma negra (7).

If repetitions of wind speeds higher than 1.0 ms^{-1} do not appear in groups, but isolated within the series, the manual inspection becomes an unaffordable procedure. In this case, a reasonable estimation of the length of a repetition considered valid is established by analyzing, first, the number of repetitions as a function of its temporal duration and second, the number of stations experiencing one or more constant records of given durations (Fig. 2.5). This second analysis was employed to assess abnormally low variations of wind direction observations by DeGaetano (1997). Both analysis show similar characteristics and either one or the other could be employed indistinctively to estimate a reasonable maximum duration of a constant wind speed episode. For short durations, constant wind periods are registered by all stations with a high frequency of occurrence. As duration increases, an abrupt decrease in the frequency of constant wind periods appears, showing that the phenomenon is less frequent. Both procedures reach zero, in absolute frequency or number of stations, at periods of eleven constant wind measurements. Constant periods of longer durations are very unfrequent in the group of stations thus suggesting eleven 10-min intervals as a reasonable choice for the maximum number of constant records allowable. Hence, constant wind speed periods equal or higher

than 1.0 ms^{-1} are considered invalid if its duration is equal or longer than 1 hour and 50 minutes. The 30-min wind speed observations only show repetitions of two records and very few of three and thus there were not corrections for them.

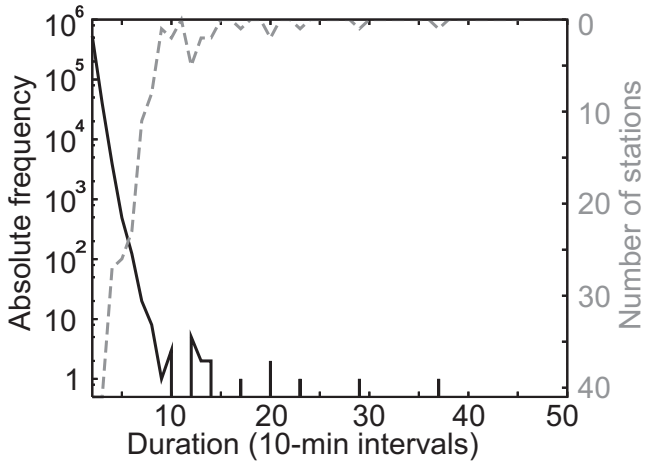

Fig. 2.5: Absolute frequency of 10-min constant periods as a function of its temporal duration (solid line) and number of stations experiencing one or more constant wind speed period of a given duration (dashed lines) for wind speeds equal or higher than 1.0 ms^{-1}. Notice the reversed scale for the number of stations.

Repetitions of wind speeds lower than 1.0 ms^{-1} are more frequent than those of higher wind speeds. This is due to the presence of calm periods which logically include the repetition of zero records and to several questionable periods that tend to report the zero value. Two of these questionable periods are displayed in Figure 2.6. The first case shows a sequence of missing observations followed by constant zero wind speed measurements during more than 20 days, and after this period, the time series recovers a realistic range of variability (Fig. 2.6a). The fact that observations within this period are preceded by missing values and the long duration of constant measurements cast serious doubts on the reliability of these data. The repetition is likely due to malfunction in the weather station.

There are some other periods of constant zero wind speed measurements that seem to be caused by freezing of the wind sensor such as that displayed in Figure 2.6b. The wind speed records show zero values for more than four consecutive days. During these days, the temperature was below 0 °C and the relative humidity close to 100 %. Under this atmospheric conditions, the excess of water vapor in the saturated air must change its phase into ice. These conditions plus a relatively weak wind, could potentially freeze the wind sensor leading to zero wind speed values. The period in which the sensor could be frozen finishes when very high wind speeds are registered. The hypothesis of freezing

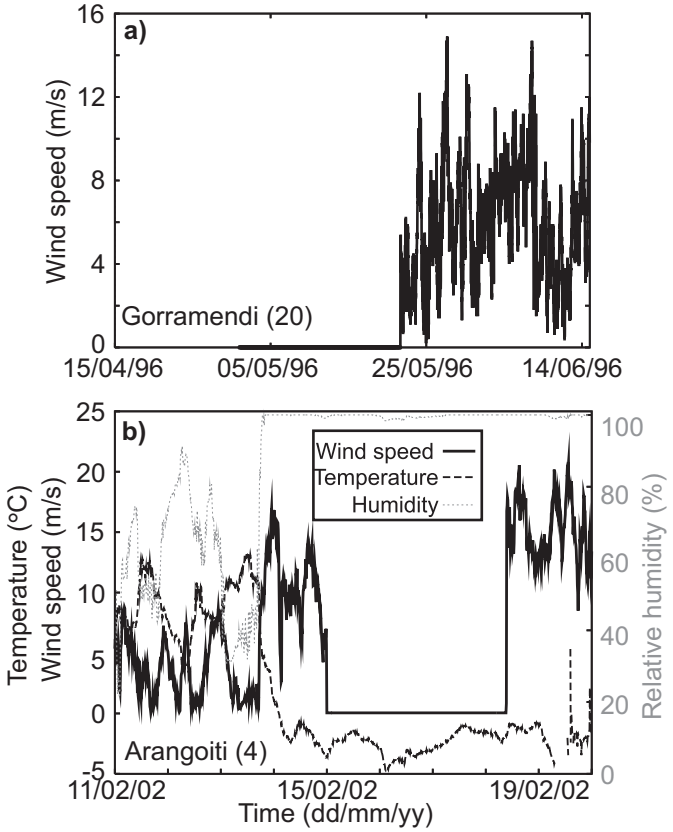

Fig. 2.6: Two different periods with constant wind speed records of 0.0 ms^{-1} from station a) Gorramendi (20) and b) Arangoiti (4). Temperature and humidity are also shown in Arangoiti for comparison.

in sensors is confirmed by metadata which reveals the presence of ice in the anemometers at the highest locations.

To suppress questionable periods like the previous ones (Fig. 2.6), a reasonable value for the maximum number of constant measurements associated with calm was estimated. This could have been done with a similar analysis used for the case of wind speeds higher than 1.0 ms^{-1} (Fig. 2.5). However, it could be argued that the durations associated with calms depend on the specific location of the station. Different sites present different wind distributions that can affect the length of the calm periods. This is illustrated in Figure 2.7a where the number of repetitions associated with calm records as a function of their duration is represented for each station. The length and frequency of calm periods depends on the location. This different behavior can be better understood comparing the number of calms with the mean wind speed. This comparison is displayed in Figure 2.7b.

The number of constant periods associated with calm decreases with the mean wind speed of the location. Therefore, the reasonable maximum duration of a calm interval should be calculated for each individual time series. This is done by analyzing the frequency of constant periods as a function of its duration for each station independently (Fig. 2.7a), in contraposition to the joint treatment of the sites for isolated wind speeds higher than 1.0 ms^{-1} (see the previous discussion concerning Fig. 2.5). The constant periods of 30-min wind speed observations lower than 1.0 ms^{-1} present a similar behavior (not shown) and thus, an equivalent analysis is applied to them.

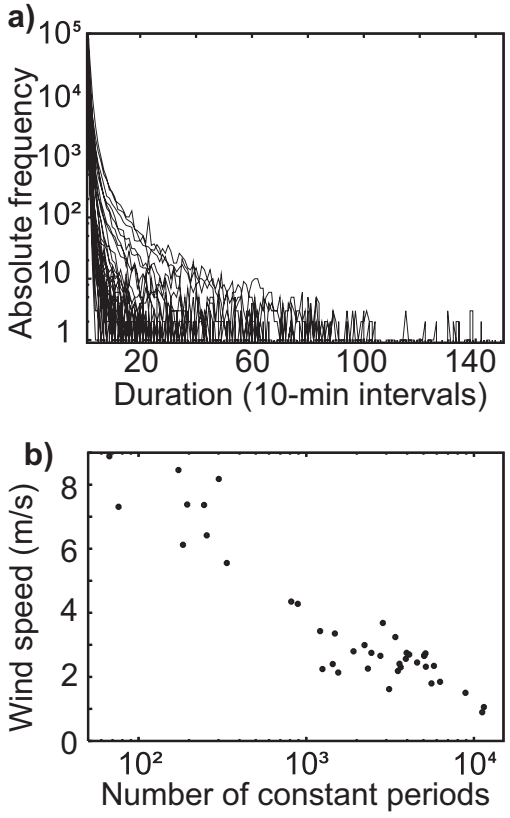

Fig. 2.7: Absolute frequency of constant periods of 10-min wind speed records for the given durations (a), and mean wind speed versus the number of constant wind speed periods. Each line (a) or point (b) represent an observational time series. Only constant periods of wind speeds lower than 1.0 ms^{-1} are considered.

After screening for abnormally low variations in the wind speed time series, the same is done for wind direction (Fig. 2.1). Zero wind direction records are excluded from the

analysis because they are associated with calm periods which were already assessed in the wind speed temporal consistency evaluation. Again, a reasonable number of admissible constant wind direction measurements is established by analyzing the frequency of constant periods of 10-min wind direction as a function of their duration and the number of stations which present one or more constant wind direction periods for the given durations. The analysis suggests forty as a reasonably limit for the maximum number of constant measurements allowed. Hence, 10-min wind direction measurements repeated during six hours and forty minutes or longer durations are considered invalid. For the 30-min wind direction measurements this occurs at four repeated records (not shown) and repetitions equal or longer than this limit, two hours, were also assumed invalid.

At this stage of the QC, a first step in the assessment of questionable periods associated with abnormally low variations from both wind speed and direction time series was accomplished (Fig. 2.1). The next step consists in a check that analyzes the frequency of constant measurements within each year before and after the corrections of the previous checks. The purpose of the analysis is to inspect the year to year variability in the number of repetitions and, if any abnormal behavior is detected, inspect the causes. The check is applied to each one of the wind speed and direction time series separately. None of the 41 stations present an abrupt decrease of the frequency of constant measurements, which would be an indicative of an excessive suppression of measurements by the previous checks. In contrast, most of the stations show a reduction in the number of repetitions in years with an abnormally high frequency of repetitions, like station Roncesvalles-Orreaga (30) where corrections suppress a peak in 1997 and other secondary peaks in 1993 and 1999, keeping the frequency of repetitions approximately constant along the nineties (Fig. 2.8a). However, during 2001 and 2002 around 80 % of the records are associated with repetitions. An inspection of the wind speed time series shows extremely low values in the period shown in Figure 2.8b. It can be noticed that wind speed records before and after this period present different variability. Due to the questinable veracity of the records, it was decided to assume invalid the period since the beginning of 2001 till the first months of 2003. Similar cases were detected by this test in other stations. All the questionable periods identified were independently analyzed in order to supresss unrealistic observations as those shown in Figure 2.8b. This check was the last one dealing with abnormally low variation of the time series (Fig. 2.1).

2.3.3.2 Abnormally high variations

The abnormally high temporal variations of wind speed records are first assessed with a step check (Fig. 2.1). This check calculates the difference between two consecutive records and if it is higher than a threshold value the second record is considered invalid. The choice of the threshold value is a controversial decision which depends on the location and the temporal resolution of the data. Meek and Hatfield (1994) recommend critical values of 7.5 ms^{-1} and 10.0 ms^{-1} for hourly and daily data, respectively. On the other hand, Bailey and McDonald (1997) suggest a value of 5.0 ms^{-1} for hourly data. Other authors (e.g. Vejen et al., 2002) propose dynamical critical values changing with the month of the year; they suggest values of 6.0, 7.0 or 8.0 ms^{-1} for hourly data at Denmark. Thus, the selection of this threshold value depends on the particular wind characteristics of the

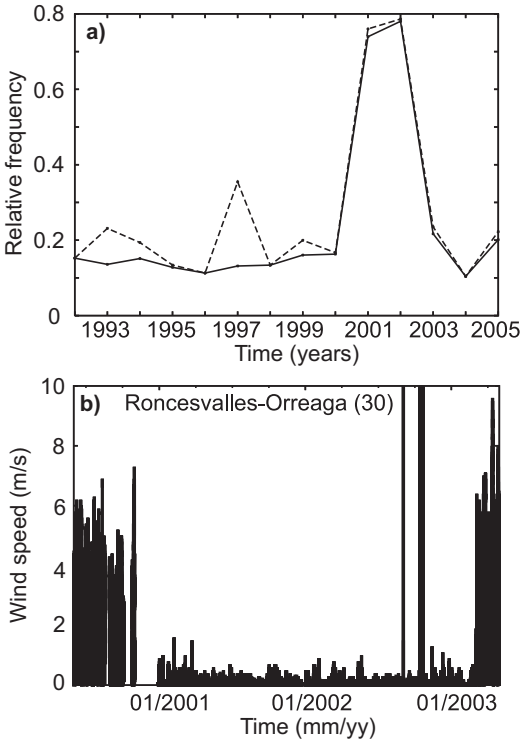

Fig. 2.8: Frequency of repeated measurements on constant records episodes during every year of the wind speed time series registered at station Roncesvalles-Orreaga (30) (a). The solid (dashed) line represents the repetitions before (after) the corrections for abnormally low variation periods. The panel b) displays a period of the wind speed time series.

region among other factors. In this QC the thresold value is selected by analyzing the histogram of the absolute differences for every wind speed time series. A representative example is displayed in Figure 2.9b. The frequency of the differences gradually decrease and it is zero at 10.0 ms^{-1}. There are two cases with larger differences of 28 ms^{-1} and 29 of ms^{-1} that cause a large break on the histogram. Actually, these two abnormally high variations are originated by the same wind speed record (Fig. 2.9a). The lenght of the break in the histogram and the low frequency of occurrence suggest that for this case 10 ms^{-1} is a good candidate for the threshold value. Other stations present a break on the histogram reaching the null frequencies at slightly different wind speeds. For instance, the stations from RN reach the null frequencies around 6 ms^{-1} for both 10-min and 30-min wind speed observational time series. These considerations lead to assume 10 ms^{-1} as a reasonable choice for the thresold value used in the step check.

The previous step check is unsuitable to assess those records which are followed or preceded by a missing value because their difference obviously cannot be calculated. In

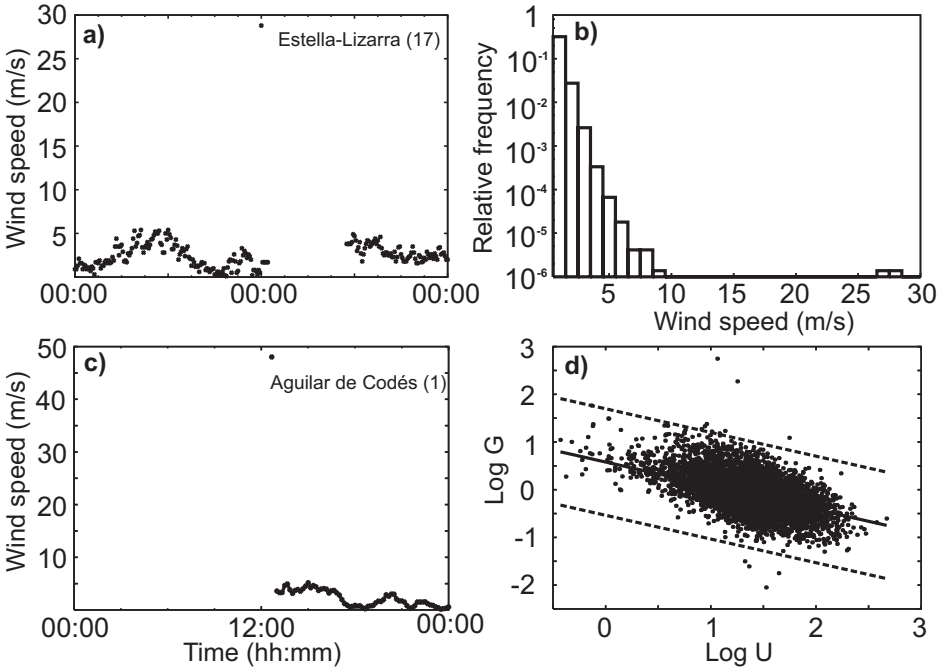

Fig. 2.9: A period of wind speed registered at a) Estella-Lizarra (17) and b) Aguilar de Codés (1). The panel b) displays the absolute value of the differences between consecutive wind speed records of Estella-Lizarra (17), and the panel d) displays the logarithm of the daily maximum 10-min wind speed normalized by the daily mean versus the logarithm of the daily mean also for station Aguilar de Codés (1). The solid line of panel d) represents the linear regression fit and the dashed lines the 95% confidence intervals.

fact, there are some abnormally high wind speed records surrounded by missing values which are not detected by the step check and should be labelled as invalid (Fig. 2.9c). In order to assess this type of records, an analysis of the daily maximum 10-min wind speed against the daily mean is performed (Fig. 2.1). This is done in a similar way as Graybeal (2006) assessed daily peak-gust records. He employed the gust factor defined as

$$G = u_g U^{-1} - 1 \tag{2.1}$$

where u_g is the peak-gust and U is the daily wind speed mean. The normalization provides a certain degree of robustness to the exposure characteristic of instrument siting (Paulsen and Schoeder, 2005). Then, he used the well-known linear relationship between logG and logU (Weggel, 1999; Jungo et al., 2002) to fit a linear regression and to establish prediction intervals. Since outliers are present in the observations, the parameters of the linear regression were calculated with a resistant technique like *the least trimmed squares regression* (Venables and Ripley, 1999). Furthermore, resistant estimators as the biweight mean and standard deviation (Lazante, 1996) were used to calculate the vari-

ance of the predictions in order to establish prediction intervals (see Graybeal, 2006, for more details). In this study, u_g is replaced by daily maximum 10-min wind speed to calculate G (Eq. 2.1). Only those days with no missing values are used to more accurately calculate the linear regression parameters and, the 95 % confidence prediction intervals, after checking that residuals are normally distributed. Once prediction intervals are calculated, all days with a minimum of ten 10-min wind speed records available are used to calculate their logG and logU. All days outside the prediction thresolds are flagged and manually analyzed to assess their questionable values. As an example, the linear fit and the prediction intervals for observations in station Aguilar de Codés (1) are displayed in Figure 2.9d. After analyzing the measurements for days outside the prediction intervals, the anomalous high value which could not be detected by the step check due to the presence of missing values is now identified (Fig. 2.9c).

The assessment of the abnormally high variations in the wind speed records completes the QC applied to the original dataset (Fig. 2.1). The QC suppresses rough errors from the dataset. It is still necessary to evaluate, and correct if possible, the potential systematic errors that could be present on the quality-controled data.

2.4 Methodologies: correction of biases

As a final step in the evaluation of the potential errors of the dataset, the systematic errors or biases of the wind speed and wind direction time series are analyzed. These errors should be detected and adjusted with a homogenization procedure (Peterson et al., 1998a). The preocedures have been mostly applied to annual and monthly averaged records (e.g. Alexanderson, 1986; Karl and Williams, 1987; González-Rouco et al., 2001). The use of the homogenization methodologies to adjust records of higher temporal resolution is delicate since the higher variability of the records introduces a new variety of problems (Aguilar et al., 2003). Recently, some effors have been oriented to develop suitable methodologies to homogenize daily temperature and precipitation time series (e.g. Vincent et al., 2002; Wijngaard et al., 2003; Della-Marta, 2006). Hence, the application of the available homogenization procedures to the high temporal resolution of the present datatset, 10-min and 30-min, is at least delicate. As a consequence, no homogenization procedure was used to adjust the potential biases of the present dataset. Instead, two checks oriented to analyze the long term behavior of the time series were applied to the data in order to deal with the most important biases. The application of these checks contribute to the homogenize the time series, but the process cannot be called in a strict sense homogenization since that would require the application of more sofisticated homogenization procedure.

The systematic errors can introduce biases in the mean or the variance of the time series. Hence, one of the checks applied deals with the long term behavior of the mean and the other one with the variance. The checks consist in the analysis of a 30 days moving average and a moving variance applied to the wind speed and direction time series. For the case of possitive defined variables, a higher mean involves higher variability (wider distribution) and thus, both statistics tend to show a direct relationship (Xoplaki et al., 2004; Jiménez et al., 2008). Thus, it can be argued that inhomogeneities on the mean

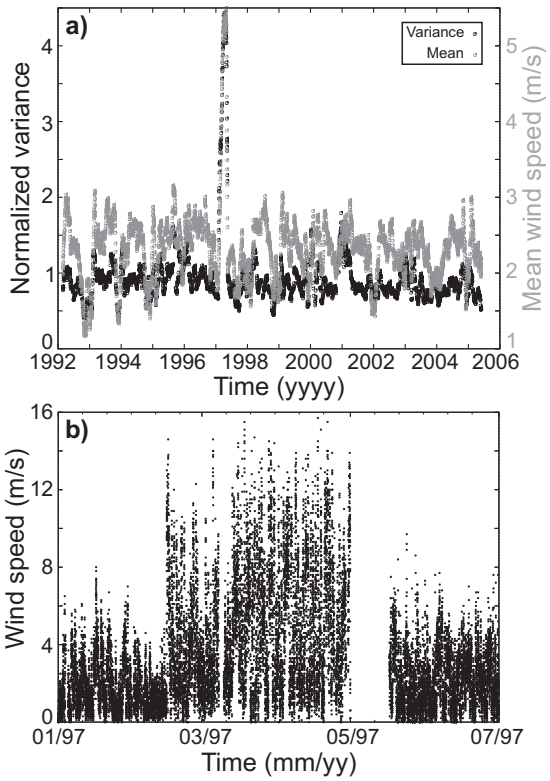

Fig. 2.10: 30 days moving variance of the wind speed records registered at station Aoitz-Agoitz (2) (a) and a selected period of the wind speed time series (b). The moving variance is normalized by the variance of the complete time series.

will also introduce inhomogeneities in the variance for this kind of variables. The wind speed is a clear example of possitive defined variable. In order to illustrate the previous theoretical considerations, the moving variance as well as the moving average of the wind speed time series registered at Aoiz-Agoitz (2) are displayed in Figure 2.10a. An excessive high increment of wind speeds variations at the begining of 1997 is recognized. The moving average shows a similar behavior. An inspecion of the records corresponding to the high values reveals a change in both mean and variance of the time series (Fig. 2.10b). After that anomalous period, a missing period appears and then, the wind speed records recover their regular range of variation. Unfortunately, no metadata is available for this period to look for possible explanations. However, the change of variance and the presence of the missing period after the abnormally high measurements weaken the reliability of this records and thus, the period was considered invalid. As in the QC, no attempt was made to fill up the missing values introduced.

The moving average and the moving variance of the wind direction time series were calculated using directional statistics (Mardia and Jupp, 1999). A common problem identified by these checks can be found in the records of station Isaba (22). The normalized moving variance oscilates around 1.0 but there are abnormally low values at the end of the time series (Fig. 2.11a). The later analysis of the questionable period shows wind direction measurements with an anomalous range of variation (Fig. 2.11b). More than one year of records are restricted to a reduced range of values (i.e. between 0° and 180°), probably due to blocks of the wind vane bearings. The complete period of anomalous variability was removed from the time series. Similar changes in variance were identified and supressed in other stations.

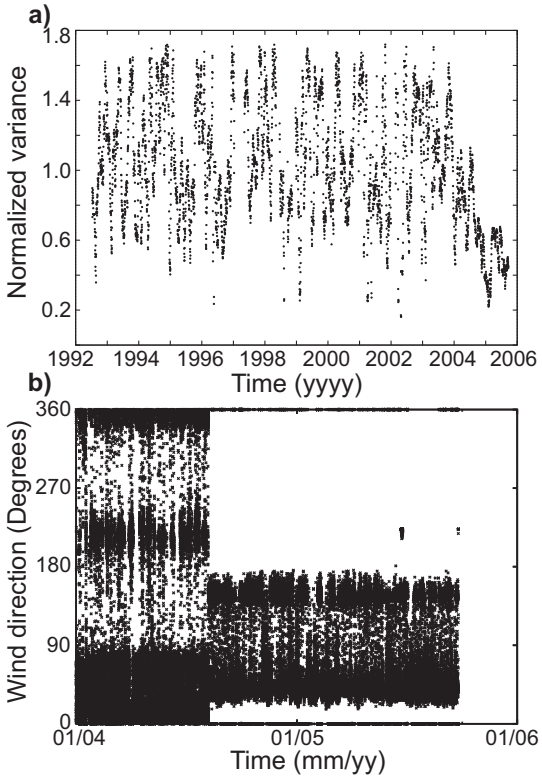

Fig. 2.11: Normalized 30 days running variance of the wind direction records registered at station Isaba (22) (a) and a selected period of the wind direction time series (b).

Other moving averages present steps in the time series associated with failures in the wind vane fixation which affect its zero reference. One of these cases can be observed in the moving average from station Pamplona-Larrabide (27) shown in Figure 2.12. It can be recognized three different steps in the moving average defining periods with different mean

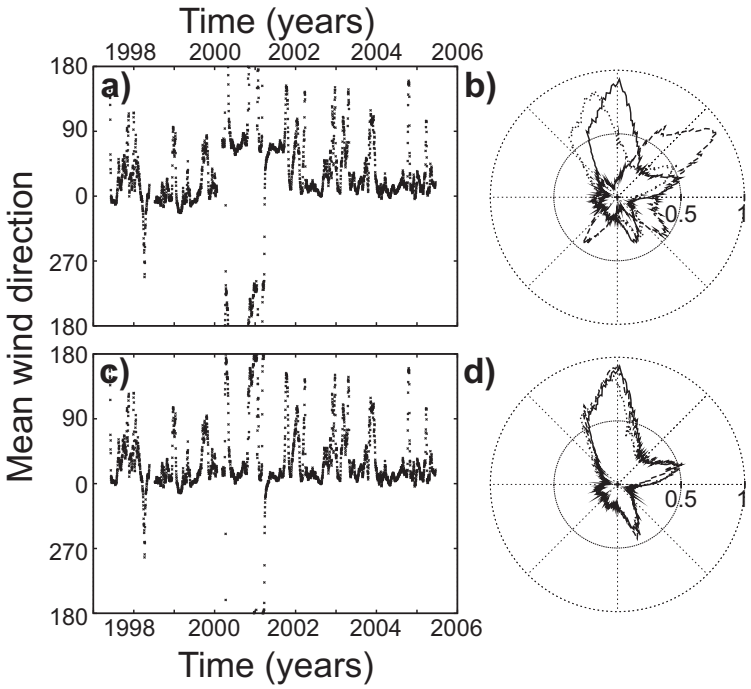

Fig. 2.12: Normalized 30 days moving variance of the wind direction records registered at station Pamplona-Larrabide (27) before (a) and after (c) the correction of its anomalous variability. The wind rose before (b) and after (d) corrections for three different periods: 1 January 1997 - 1 February 2000 (dotted line), 1 February 2000 - 18 September 2001 (dashed line) and 18 September 2001 - 27 June 2005 (solid line) is also shown.

(Fig. 2.12a). Inspection of the records in each period does not reveal anything suspicious. However, metadata show a wrong setting of the reference (North) in 18 September 2001, suggesting that a correction of 60° should be applied to the previous measurements. This date coincides with the abrupt change observed in the moving average time series which defines the end of the second period shown in Figure 2.12a. The wind roses for the mentioned periods present similar shape but they are rotated with respect to each other (Fig. 2.12b). The wind rose of the second period needs to be rotated 57° to reach a good concordance with the one corresponding to the third period (Fig. 2.12d). This is in good agreement with metadata suggestions and therefore 57° are subtracted to records of the second period. Furthermore, knowing that the wind vane was fixed in a certain date defines a reference for a good calibration of measurements and, 10° is the correction applied to the first period to be in concordance with the rest (Fig. 2.12d). After these corrections, the moving average of wind direction records does not present abrupt changes and therefore the time series presents similar mean wind direction (Fig.

2.12c). All the questionable periods showing this kind of problem were manually anlyzed in order to compensate the systematic errors of the time series.

2.5 Results: quality control

This section presents the spatio-temporal distribution of records removed by each check of the QC (Fig. 2.1). The number of measurements invalidated by each check of the QC and the correction of biases are summarized in Table 2.2.

Table 2.2: Number of invalidated records by each step of the QC (Fig. 2.1) and by the correction of biases. The percentage of invalidated records with respect to the original number of records is displayed in brackets (%).

Check	Wind speed	Wind direction
Manipulation error (temporal)	23128 (0.120)	22810 (0.118)
Manipulation error (spatial)	9909 (0.051)	9835 (0.051)
Range check	518 (0.003)	555 (0.003)
Abnormally low variability		
High wind speeds (groups rep.)	51592 (0.267)	———
High wind speeds	375 (0.002)	———
Low wind speeds	119112 (0.616)	———
Wind direction	———	96796 (0.501)
Repetitions along years	133587 (0.691)	92624 (0.479)
Abnormally high variability		
Step check	305 (0.002)	———
Daily mean vs daily max 10-min	719 (0.004)	———
Biases	10450 (0.054)	497945 (2.576)
Total	349695 (1.809)	720565 (3.728)

2.5.1 Manipulation errors

The first step of the QC screens for errors associated with the management and storage of the data (Fig. 2.1). A total of 33037 wind speed records (0.171 %) and 32645 wind direction records (0.169 %) were removed from the original time series in this step. Errors associated with a temporal misplacement of observations are more frequent than those associated with a wrong spatial assignment (Table 2.2). The spatial and temporal distributions of the records assumed as invalid by each one of the two checks are displayed in Figure 2.13.

Stations from the three institutions GN, AEMET and RN present coincident observations in their own time series here attributed to a wrong temporal assignment of observations (Fig. 2.13a). The number of cases assumed as invalid shows a maximum in 1999 (Fig. 2.13b). During this year, the wind speed and direction records for 1 November

and 1 December were coincident for a total of nine stations, all from GN. The magnitude of the wind speed measurements shows abrupt breaks before and after the coincident records for the time series that include the 1^{st} of November (not shown). On the contrary, the time series that include the records associated with 1 December do not show abrupt changes at these points. This different behavior seems to suggest a wrong assignment from the observations from the 1^{st} of December 1999 to the 1^{st} of November 1999 that contribute to the high number of cases detected during this year (Fig. 2.13b).

Most of the records assumed as invalid by the spatial check are due to coincident observations between stations managed by GN (Fig. 2.13a). The only exception is one station belonging to AEMET. The total number of records assumed as invalid by the two checks show a tendency to decrease along years (gray line in Fig. 2.13b).

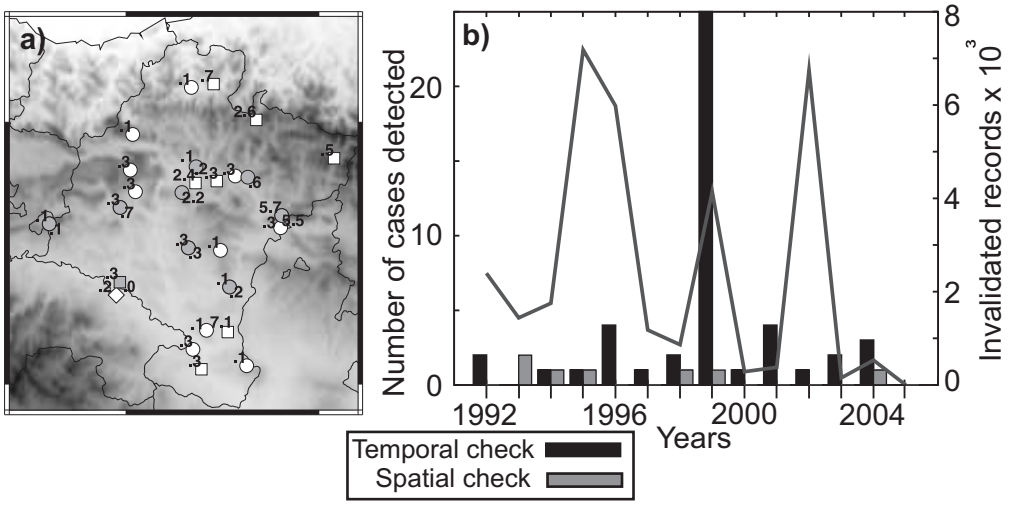

Fig. 2.13: Spatial (a) and temporal (b) distributions of records invalidated by the checks associated with manipulation errors (Fig. 2.1). Circles, squares and diamonds in panel a) represent the stations of the GN, AEMET and RN, respectively. The numbers to the left (right) of the symbols denote the thousands of records invalidated by the temporal (spatial) check. The symbols of the stations that at least contain corrections associated with the spatial check are filled up in gray. The black (gray) boxes in panel b) represent the number of periods with coincident records assumed as invalid by the temporal (spatial) check whereas the gray line represented the thousands of records invalidated by both manipulation error checks.

2.5.2 Limits consistency

This check ensures that observations of the different variables are between the minimum and maximum allowed values (see Section 2.3.2). There were not negative wind speed or

wind direction records in the observational time series but, there were 518 wind speed measurements (0.003 %) higher than the maximum value measurable by the anemometers and 555 wind direction records (0.003 %) higher than 360° (Table 2.2). The spatial and temporal distribution of the records removed after the application of the range check are shown in Figure 2.14.

The highest number of wind speed records assumed as invalid come from station Arangoiti (4), that presented 429 consecutive records of 444.4 ms^{-1} (Fig. 2.14a). The rest of wind speed records invalidated are isolated values in the time series of stations managed by AEMET. For the wind direction, the supressed records appear close to each other defining an anomalous period. Only three periods were identified, the wind direction records associated with the previous records of 444.4 ms^{-1} that show values of 444.4 degrees, and two periods registed at stations Aralar (3) and Arangoiti (4, Fig. 2.3). The invalidated records from both wind speed and wind direction time series peak in years 1999, 2000 and 2001 revealing an inhomogeneous ocurrence of these kind of errors in the dataset (Fig. 2.14b).

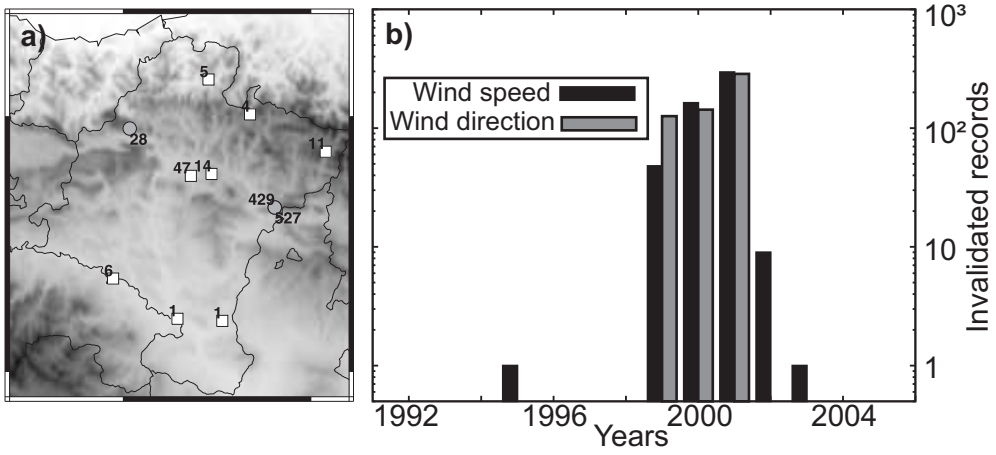

Fig. 2.14: Spatial (a) and temporal (b) distributions of records invalidated by the range check. Circles and squares in panel a) represent the stations of the GN and AEMET, respectively; the numbers to the left (right) of the symbols denote the number of wind speed (direction) records invalidated. The symbols of the stations that at least contain some suppressions of wind direction records are filled up in gray. The black (gray) boxes in panel b) represent the number of wind speed (direction) records invalidated along the years.

2.5.3 Temporal consistency

Wind speed and wind direction present distinct physical properties and thus, the procedures applied to ensure the temporal consistecy of the observational time series are

different for each variable (Fig. 2.1). The next two sections summarize the results for each of them.

2.5.3.1 Wind speed

The first step to assess the temporal consistency of wind speed time series deals with periods with abnormally low variations (Fig. 2.1). A total of 304666 wind speed records (1.576 %) were invalidated by this group of checks. The spatial and temporal distributions of the measurements assumed invalid by each test are displayed in Figure 2.15.

The supression of groups of repetitions of wind speed higher than 1.0 ms^{-1} removes 51592 records (0.267 %) from the original dataset (Table 2.2). All groups of repetitions are similar to the one displayed in Figure 2.4, and all the stations affected belong to GN (Fig. 2.15a). The single repetitions of wind speed higher than 1.0 ms^{-1} assumed as invalid are less frequent. Only 375 records were removed (0.002 %). The records invalidated by this and the previous check are concentrated in the 1993 to 1997 period (Fig. 2.15b), suggesting that the causes that originated these questionable records were fixed.

A total of 119112 (0.616 %) records were supressed for the case of repetitions of wind speeds lower than 1.0 ms^{-1}. A high percentage of the invalidated records is associated with long periods with constant zero records such as the one displayed in Figure 2.6a. The stations managed by AEMET and the stations managed by GN located at mountain tops (3, 4, 16, 20, 35 and 37, see Fig. 1.2) show the highest number of records suppressed (Fig. 2.15c). As a result, the stations managed by GN and AEMET show the largest contributions to the total number of records invalidated along the years (Fig. 2.15d). It is interesting to notice that the 9 stations managed by AEMET present the same order of magnitude in the number of invalidated records as GN, that manages 25 stations. The 7 stations managed by RN show less frequenctly this kind of problematic records.

The check that evaluates the frequency of repetitions along the years invalidates a high percentage of measurements (0.691 %) from the wind speed time series (Table 2.2). This is due to the presence of long questionable periods such as the one displayed in Figure 2.9b. The stations that showed the largest number of corrections are again some stations managed by AEMET and the stations managed by GN located at the mountation tops (Fig. 2.15e). The distribution of the anomoalous periods detected by this check is broadly homogeneous through the years (Fig. 2.15e).

All the previous analyses assess the abnormally low variations of wind speed observations in a first step to ensure the temporal consistency of the series. In a sencond step, the abnormally high variations are analyzed (Fig. 2.1). The step check invalidates 305 wind speed records (0.002 %) and the analysis of the daily maximun wind speed versus the daily mean 719 (0.004 %). The spatial and temporal distribution of the invalidated measurements is displayed in Figure 2.16.

Again, it can be appreciated that certain stations from AEMET and the stations from GN located at mountain tops accumulate the lagest number of removed records (Fig. 2.16a). The number of measurements flagged shows a homogeneous distribution through the years for both checks (Fig. 2.16b). The surpressions affect few enough wind speed observations to assume them invalid without disturbing the dataset characteristics.

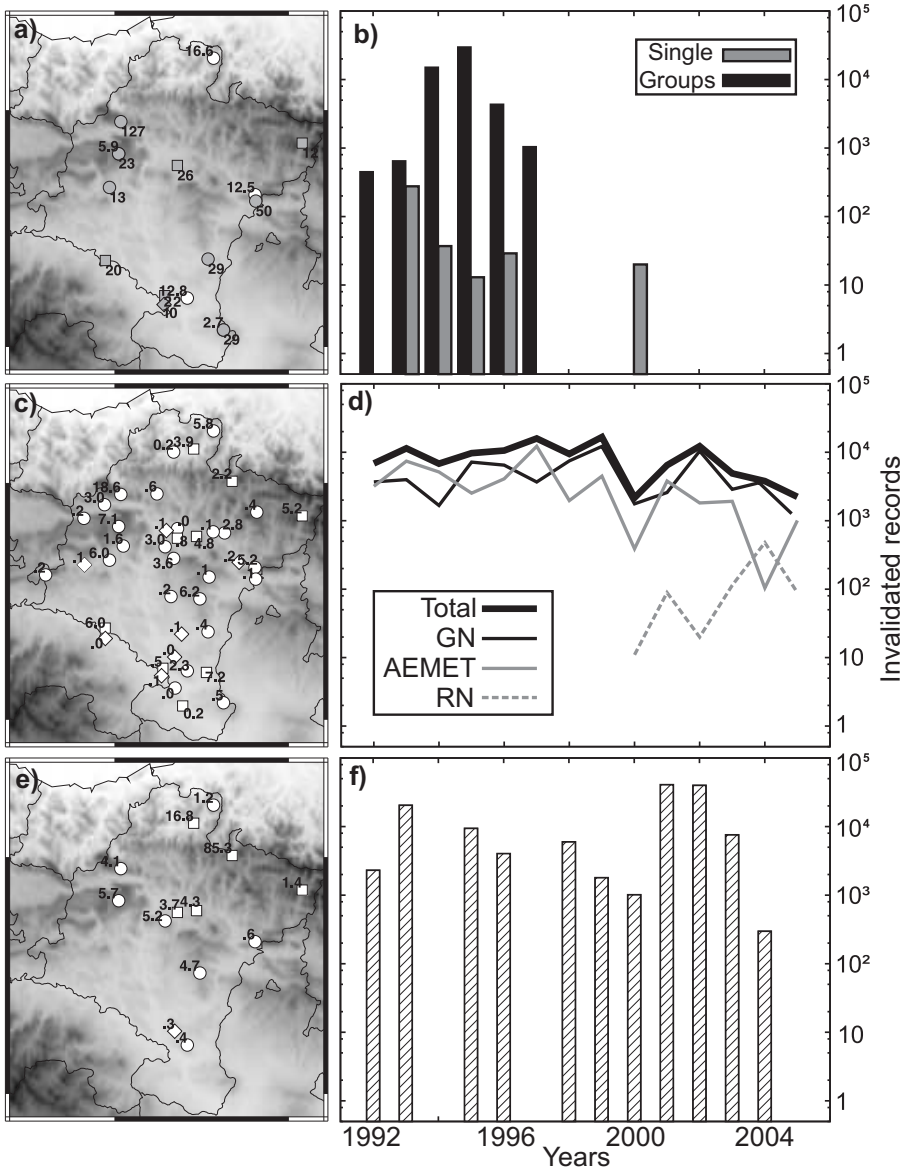

Fig. 2.15: The numbers to the left (right) of the symbols in panel a) denote the thousands (number) of records invalidated by the check that deals with groups of (single) repetitions of wind speeds higher than 1.0 ms^{-1}. The black (gray) boxes in panel b) represent the number of records assumed as invalid by the check dealing with groups of (single) repetitions. Panel c) shows the thousands of records assumed invalid after suppressing questionable periods of wind speeds lower than 1.0 ms^{-1}; and the panel d) represents the number of invalidated records form the stations of each institution along the years. Panel e) shows the number of records invalidated after analyzing the year to year variability in the number of repetitions and panel f) shows the temporal distribution of these suppressions.

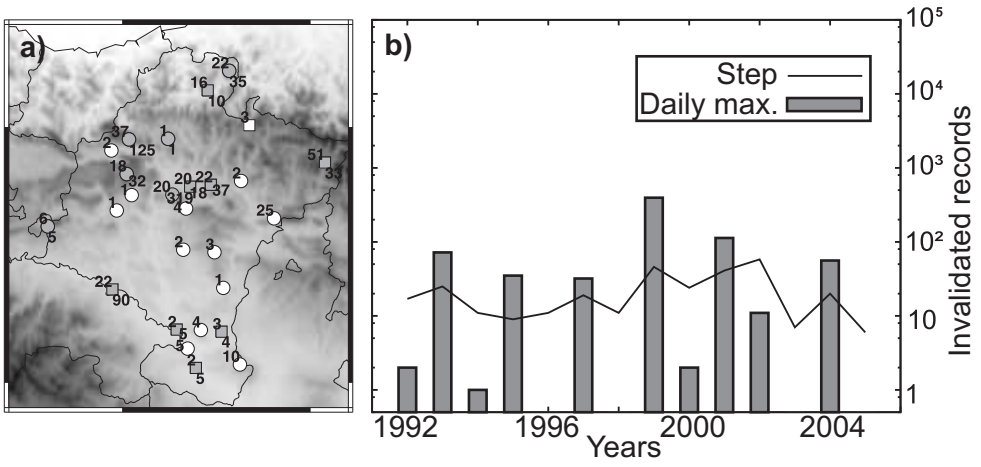

Fig. 2.16: Spatial (a) and temporal (b) distributions of records invalidated by the checks that deal with the abnormally high variations in the wind speed time series (Fig. 2.1). Circles and squares in panel a) represent the stations of the GN and AEMET, respectively. The numbers to the left (right) of the symbols denote the number of wind speed records invalidated by the step (daily maximum wind speed) check. The symbols of the stations that at least contain some suppressions by the analysis of the daily maximum wind speed are filled up in gray. The solid line (gray boxes) in panel b) represent the number of wind speed of records invalidated by the step (daily maximum) check.

2.5.3.2 Wind direction

The temporal consistency of wind direction records is ensured by removing abnormally low variations from the time series (Fig. 2.1). A total of 189490 wind direction records (0.980 %) were assumed invalid (Table 2.2). The spatial and temporal distributions of the measurements assumed invalid by each test are displayed in Figure 2.17.

The check that removes isolated repetitions considers invalid a total of 96796 (0.501 %) wind direction records (Table 2.2). Stations managed by the three institutions, GN, AEMET and RN were affected by these corrections (Fig. 2.17a). The check that analyses the frequency of constant records in every year eliminates 92624 (0.479 %) additional records (Table 2.2). The records invalidated by these checks seem to be more frequent in recent years (Fig. 2.17b), probably due to the larger number of operative stations at the end of the observational period (Table 2.2). Most invalidated observations are constant or nearly constant measurements during a long period. For instance, station Bardenas-Barranco salado (6) showed repetitions around 360° during more than one year just after the change in temporal resolution of its records (from 10-min to 30-min), suggesting a failure in the adjustments performed. The first check for isolated repetitions partially supressed the erroneous records, 15.9 thousands (Fig. 2.17a). The rest of records, 11.3 thousands, were invalidated by the check analyzing the frequency of repetitions along the

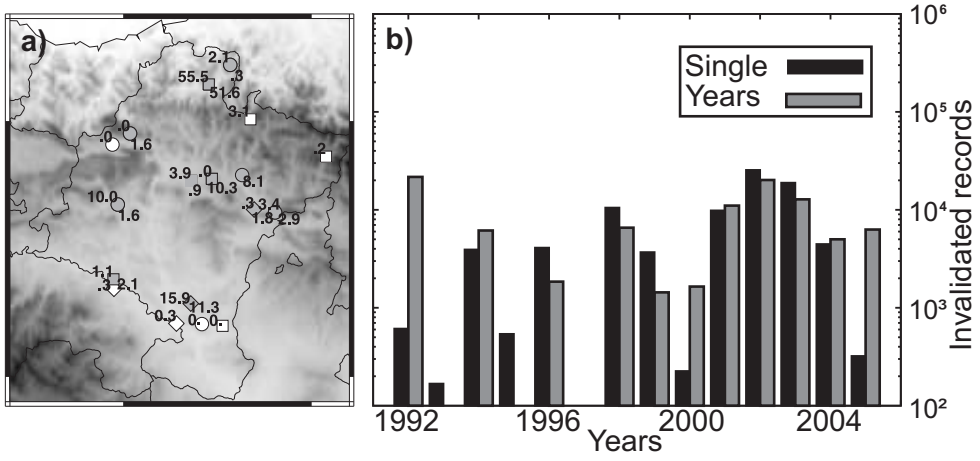

Fig. 2.17: Spatial (a) and temporal (b) distribution of records invalidated by the checks associated with abnormally low variations of the wind direction time series (Fig. 2.1). The numbers to the left (right) of the symbols in panel a) denote the thousands of records invalidated by the check that deals with single repetitions (frequency of repetitions along the years). The symbols of the stations that at least contain corrections associated with the analysis of repetitions along the years are filled up in gray. The black (gray) boxes in panel b) represent the number of records assumed as invalid by the check dealing with single repetitions (repetitions along years).

years. A similar period was registered at Lekaroz (23) that also shows a large number of invalidated records by these checks (Fig. 2.17a).

The inordinately long length of some of the invalidated periods produce large impacts on the statistics of the wind direction time series. For instance, the mean direction registered in Lekaroz (23) shows a change of 19 degrees after the application of the check dealing with isolated repetitions to the time series. The supressions introduced after the analysis of repetitions through years produce an additional change of 13 degrees. The variance shows a weaker impact with changes of about 1 % for each check. Station Bardenas-Barranco salado (6) shows changes in the mean of 6 and 7 degrees after the application of each ceck. The standard deviation also shows a weaker impact with changes of about 5 %.

2.6 Results: correction of biases

This section summarizes the results of applying the checks oriented to analyze the systematic erros.

The inspection of the 30 days moving means (variances) of the wind speed time series does not reveal suspicious steps. There exist several abnormally high peaks but they present spatial consistency. The only case finally removed was the questionable period registered at Aoiz-Agoitz (2) displayed in Figure 2.10. This correction removes 10450

(0.054 %) wind speed records and causes a decrease in the mean wind speed (variance) of 0.04 ms^{-1} (4%).

The screening for biases on the wind direction time series helped to identify periods with abnormal ranges of variation (e.g. Fig. 2.11) and other periods associated with a wrong setting of the wind vane (e.g. Fig. 2.12). The errors associated with an anomalous range of variation usually present a long duration and thus, a high percentage of records (2.576 %) are invalidated by this check (Table 2.2). Problems with the fixation of the wind vane affect to a total of thirteen stations. The spatial and temporal distribution of these corrections are displayed in Figure 2.18.

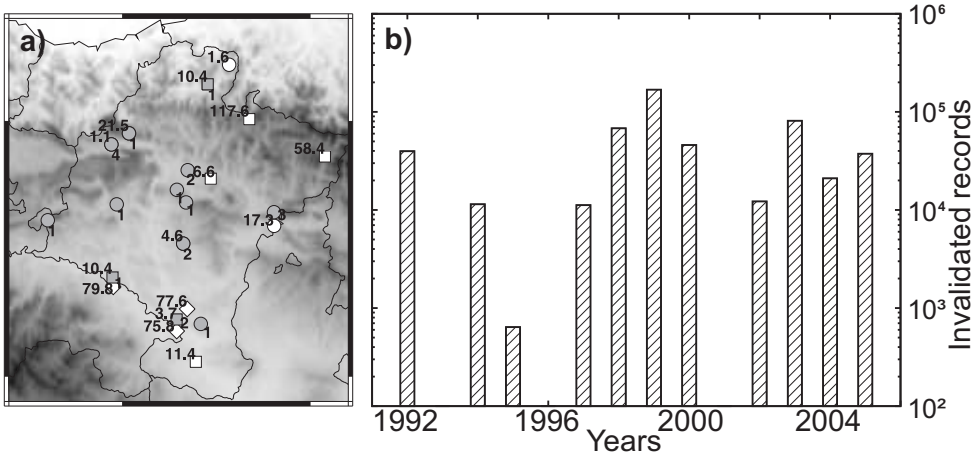

Fig. 2.18: Spatial (a) and temporal (b) distribution of records invalidated by the checks associated with biases in the wind direction time series. The numbers to the left (right) of the symbols in panel a) denote the thousands of records invalidated (number of periods rotated). The symbols of the stations for which records were rotated at least in one temporal period are filled up in gray.

Stations from the three institutions register anomalous periods of variation that affect a large number of records (Fig. 2.18a). The temporal distribution of the invalidated records reveal a rather continuous occurrence of this problem along the years (Fig. 2.18b). The problems with the fixation of the wind vane are corrected in a similar way as in station Pamplona-Larrabide (27) when metadata are available (Fig. 2.12). Otherwise, measurements corresponding to different periods of variability are rotated to be in concordance with the last period. This ensures that future extensions of the wind dataset will present similar wind direction variability. The only exception was station Cadreita-INM (12) in which no metadata were available and the different periods were rotated to be in concordance with the wind rose from station Cadreita-Riegos (11), a very close station (Fig. 1.2) that does not present anomalous wind direction periods.

2.7 Impacts

This section describes the impacts that the QC and the correction of biases introduced in the statistics of the observational time series. The percentage of records removed from each one of the original time series are displayed in Figure 2.19. The percentage of missing values introduced in the wind speed time series is lower than 5 % except for station Roncesvalles-Orreaga (30) that shows a percentage higher than 10 % (Fig. 2.19a). The large number of invalidated records at this site is a result of suppressing the long questionable period displayed in Figure 2.8b by the check dealing with the frequency of repetitions over the years. The missing values introduced in the time series of the wind direction are also in general lower than 5 % (Fig. 2.19b). However, a total of six stations show higher percentages of missing values which in five cases exced 10 % (Fig. 2.19b). These large percentages are associated with the checks dealing with the repetitions over the years as well as corrections of biases.

The impacts that the supressions produced in the mean of the time series are displayed in Figures 2.19c and 2.19d for the wind speed and wind direction, respectively. Changes in the mean wind speed show a maximum value of 0.3 ms^{-1} at the station Roncesvalles-Orreaga (30). The wind direction time series show larger impacts with changes in the mean direction higher than 30°. The standard deviation of the wind speed time series decreased in all stations (Fig. 2.19e) due to the invalidation of outliers and abnormally long periods recording zero wind records. The standard deviation of the wind direction time series shows moderate changes with the largest one at Bardenas-Barranco Salado (6, Fig. 2.19f) due to the supression of the long period with repetitions around 360° already mentioned.

For the case of the wind speed time series, it is interesting to analyze the impacts that the corrections produced in higher order moments suchs as the skewness and the kurtosis (Fig. 2.20). The high values that the original time series showed at some locations (Fig. 2.20a,c) are reduced after the corrections (Fig. 2.20b,d). The skewness is positive at all locations revealing that the tail of the wind speed distributions is larger in the positive direction. The kurtosis also tends to be positive, indicating a higher peak than a normal distribution at most of the sites.

The 10-min and 30-min quality-controlled observations at each station were daily averaged for the specific purposes of this study. Only those days with more than 50 % of the records available were used. The rest of days were flagged as missings. In addition, those days lacking either the wind speed or the wind direction were also flagged as invalid. The spatial and temporal distribution of days with available daily mean wind is displayed in Figure 2.21. Stations showing a large percentage of missing values are those locations in which the sensor was installed more recently (see Table 2.1). This causes an hetereogeneous distribution of the number of available observations over time (Fig. 2.21b). There are approximately 20 stations with observations available prior to 1999, when the number of stations starts to increase. The drop in the number of observations available in 2005 is a consequence of a sequential delivery of the wind observations. Futures updates of the dataset will solve this limitation.

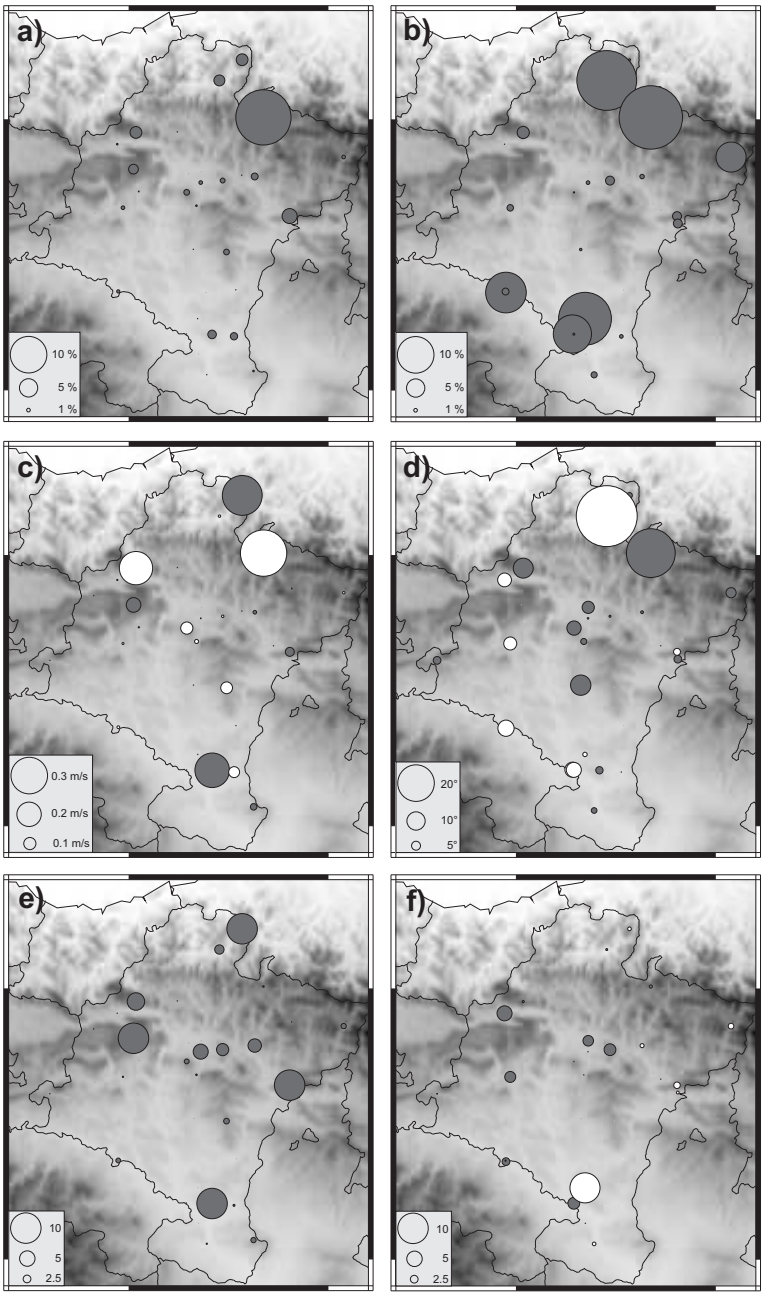

Fig. 2.19: Percentage of records removed from the original wind speed (a) and direction (b) time series. The bias (original minus final) is displayed in panel c) for the wind speed and in panel d) for the wind direction. A negative (positive) bias is displayed in white (black). Panels e) and f) show the ratio between the original and final standard deviations for the wind speed and wind direction time series, respectively. A ratio lower (higher) than one is displayed in white (black).

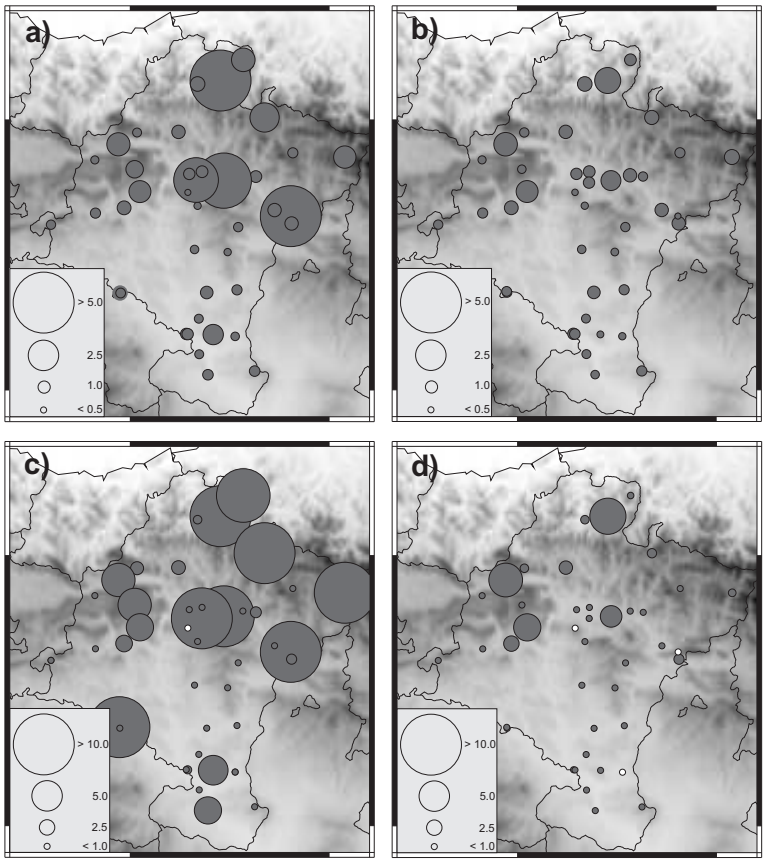

Fig. 2.20: Skewness of the wind speed time series before (a) and after (b) the QC and the correction of biases. The kurtosis before and after the supressions is displayed in panels c) and d), respectively. Positive values are displayed in black and negative ones in white.

2.8 Conclusions

The potential errors of the wind speed and wind direction observations acquired at the 41 automatic weather stations over the CFN used in this Thesis have been analyzed. Rough errors were removed with a QC that screens the observations for 1) manipulation errors, 2) limits consistency and 3) temporal consistency (Fig. 2.1). Then, the systematic errors were analyzed and corrections derived when possible. About 1.8 % wind speed and 3.7 % wind direction records were assumed invalid (Table 2.2). The abnormally low variation is the principal cause of wind speed records invalidation. On the other hand, systematic errors associated with an anomalous range of variation are the most important source of invalid wind direction measurements. All the problems that have been detected and

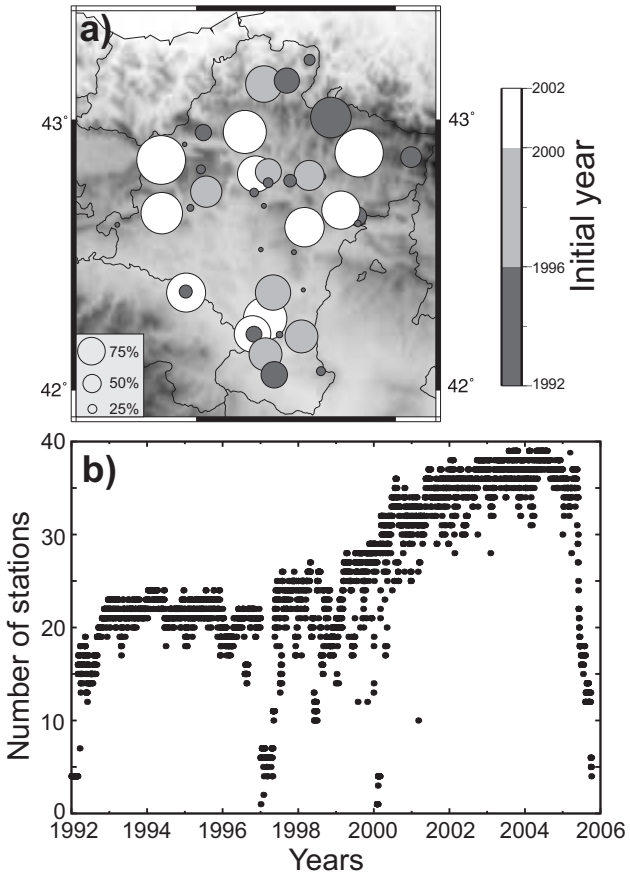

Fig. 2.21: Percentage of days (circles) with missing daily wind (a). The color of the circle displays the first year with an observation available. Panel b) shows the time series of the number of stations with the daily mean wind available.

particularly the ones that persist during a certain time, should be considered to improve the quality of the observations in the future.

The corrections applied have been shown to have a clear impact on the general statistics of the dataset. The supression of these records from the dataset is also very important for the subsequent analyses of data and model evaluation where dealing with erroneous data in specific periods can lead to missleading conclusions. It should be noticed that a small number of corrections in a time series means a good quailty of the observations, which could not have been guaranteed before the application of the present depuration process. In this sense, the QC and correction of biases process confer more robustness to the conclusions reached in the following chapters.

It should also be pointed that the quality controlled data is not expected to be an error-free dataset, but is one where the quality of the time series has been significantly improved. Future updates on the dataset and the testing of new methodologies will likely lead to new quality corrections.

3

Dynamical downscaling: the WRF numerical simulation

The objective of this project is an investigation of the theory of dynamic meteorology in order to make it accessible to high-speed, electronic, digital, automatic computing, of a type which is beginning to be available, and which is likely to be increasingly available in the future. ... Indeed, the possibilities that are opened up by these devices are so radically new and unexpected, that the theory is entirely unprepared for them. There was no practical motivation in the past to work out those parts of meteorologcial theory on a mathematical and analytical level, which in order to become really effective, would require calculations faster than what seemed possible at the time! A complete reassessment or reevaluation of the theory is therefore an absolute prerequisite.

J. von Neumann, 1946.

The previous Chapter 2 described the observational dataset used in this study and the procedures applied to improve its quality. The data will be used to analyze the wind behavior and to evaluate the capability of a dynamical downscaling approach to reproduce it. This Chapter 3 describes the general characteristics of the WRF model used to perform the downscaling, and its particular configuration for this study. The comparison of the simulation with observations will be presented in Chapters 5 and 7.

3.1 A hierarchy of models available to estimate the wind field

Although the CFN observational network (Fig. 2.1) offers a good spatial coverage of the region, one would always like to increase the number of observational sites. Of course, the mind of a researcher always flies further than what can be achieved in reality. But even if it would be possible to make such dreams reality, there is no possibility to recover lost information: failures in the observational system, a late installation of the sensor, etc. No matter how accurately one defines his/her "perfect" observational network, there will be always a *but*. In the absence of enough observations to provide the desired representation of the regional flow, which unfortunately is too often the case, the available information can be used to estimate the wind at other locations.

The wind estimations should be based on the physical laws governing the atmospheric evolution. To our current knowledge this is represented by the Navier-Sotokes equations, or the Euler equations if it is assumed that the atmosphere is an ideal fluid. Both systems of equations, in a strict sense, only refer to the conservation of momentum. It is necessary to introduce the conservation of energy and mass as well as a equation of state (the ideal gas law) to close the system of equations. The estimated wind field needs to satisfy totally or partially, depending on the assumptions and simplifications, these equations.

Perhaps the simplest estimations (if there is anything simple) are obtained by means of a diagnostic model. These models blend observational information and modeling to provide a physically consistent interpolation of the wind field. This is achieved by imposing constraints as in the case of the mass-consistent models proposed by Sherman (1978) that impose the conservation of mass through the continuity equation (e.g. Sherman, 1978; Goodin et al., 1980; Endlich et al., 1982). Somewhat more elaborated diagnostic models are based on the theory of Jackson and Hunt (1975) and its extension to three dimensions by Mason and Sykes (1979) that linearized the equations of motion to obtain an analytical solution (e.g. Walmsley et al., 1982; Troen and Petersen, 1989). These models provide satisfactory results over hilly terrain (Jenkins et al., 1981; Mason and King, 1984) but their application to steep slopes typical of complex terrain regions can be problematic due to the linearization applied. In addition, diagnostic models do not take into account thermal effects and thus, their use is restricted to evaluate the effects of orography on steady mean wind flows (Ratto et al., 1994).

A more realistic representation of the physical processes which include thermal effects is provided by prognostic mesoscale models (Pielke, 2002). These models numerically solve the Euler equations, usually after applying further simplifications such as the Reynolds averaging (Reynolds, 1895), to predict the meteorological fields (e. g. Black, 1994; Grell et al., 1994; Cotton et al., 2003; Skamarock et al., 2005). Mesoscale models have become a standard tool to provide simulations and forecasts of air flows in complex terrain regions, favored by recent increases in computational power and accessibility of analysis and forecast grids (Mass and Kuo, 1998).

The WRF mesoscale model (Skamarock et al., 2005) is used in this work to increase the spatial and temporal sampling of the observational network. In order to provide a good temporal coverage and a reasonably accurate reproduction of the surface heterogeneity, the simulation spans the complete observational period (1992-2005) at a high horizontal resolution (2 km) over the CFN. Of course, a wind estimation is just that, an estimation, and it would be of no practical utility, no matter how elaborate is the model used to derive it, if it is not compared against observations to confirm its agreement. It is not necessary to say the implications that an accurate enough wind estimation can present for quite different issues.

3.2 An overview of the advanced research WRF

The advanced research WRF version 2.1.2 is used to perform the numerical simulation. The information here provided follows Skamarock et al. (2005) where the interested reader is referred to for a more complete description of the model.

The numerical algorithm integrates the compressible, nonhydrostatic Euler equations following the philosophy of Ooyama (1990). The equations are formulated using a η terrain-following hydrostatic-pressure vertical coordinate (Laprise, 1992) whereas the horizontal discretization is a C-grid staggering (Arakawa and Lamb, 1977). The basic prognostic variables are the velocity components, the air potential temperature, the mixing ratios of hydrometeors, and the perturbations from a hydrostatic reference state of the geopotential and surface pressure of dry air.

A time-split integration scheme is used in the temporal discretization (Skamarock and Klemp, 1992; Wicker and Skamarock, 2002). The low frequency modes are integrated using a third-order Runge-Kutta time integration scheme. The high-frequency acoustic modes are integrated over smaller time steps to maintain numerical stability. The horizontally propagating acoustic modes are integrated using a forward-backward time integration scheme, and the vertically propagating acoustic modes and buoyancy oscillations are integrated using a vertically implicit scheme. The time-split integration is similar to that first developed by Klemp and Wihelmson (1978).

A set of packages is provided to take account for physical processes which are not explicitly solved in the equations. These packages or parameterizations estimate the effects associated with mycrophysics, cumulus, planetary boundary layer (PBL), surface layer, land surface, radiation, and horizontal diffusion.

Mycrophysics includes explicitly resolved water vapor, cloud and precipitation processes whereas the cumulus parameterizations are responsible for the sub-grid-scale effects of convective and/or shallow clouds. The cumulus schemes simulate the effects of vertical fluxes due to unresolved updrafts and downdrafts and compensating motion outside the clouds. These parameterizations infer the presence of clouds on each vertical column, and provide estimations of the vertical heating and moistening profiles produced by the clouds as well as the convective component of the surface rainfall.

The surface layer schemes calculate friction velocities and exchange coefficients that enable the calculation of surface heat and moisture fluxes by the land-surface models and surface stress in the PBL scheme. The PBL is also responsible for vertical sub-grid-scale fluxes due to eddy transports providing atmospheric tendencies of temperature, moisture, and horizontal momentum in the whole atmospheric column (not just the boundary layer). The horizontal dissipation/diffusion schemes take account for the horizontal turbulent mixing. The physical packages of the surface layer, PBL and land-surface model are one-dimensional column schemes.

Finally, the radiation schemes provide heating/cooling due to radiative flux divergence and surface downward and shortwave radiation for the ground heat budget. Longwave radiation includes infrared radiation absorbed and emitted by gases and surfaces. Shortwave radiation includes visible and surrounding wavelengths whose only source is the Sun, but includes the absorption, reflection, and scattering in the atmosphere and at surfaces. The available radiation schemes are also on a one-dimensional column.

A more detailed explanation of the parameterizations can be found in Stensrud (2007).

3.3 Model configuration

Dynamical downscaling exercises frequently use reanlysis data as initial and boundary conditions since they represent our most accurate knowledge of the atmospheric states. This is the case of the WRF simulation performed in this Thesis. The available reanalyses typically present and horizontal resolution of about one degree (Kalnay et al., 1996; Uppala et al., 2005), and this characteristic strongly conditions the design of a numerical experiment. The selection of the data used to drive the downscaling is an important decission, but the design of the experiment also requires a careful model configuration to analyze the potential implications that the dynamical and physical selections can introduce on the simulation.

3.3.1 Dynamical configuration

The size of the area over which the atmospheric evolution is simulated emerges as a first important setting (Vannttsem and Chomé, 2005; Rauscher et al., 2006). On the one hand, it could be argued that a large area would be desirable in order to locate the target region, the CFN in this study, far from the boundaries and thus reduce the influences of the error propagated from the lateral boundary conditions (Warner et al., 1997). A large domain would be therefore beneficial to allow the model to develop its own internal variability. On the other hand, it could be argued that a smaller domain size constrains the atmospheric evolution towards the one imposed on the boundaries, which could be desirable if the forcings are derived from observations (Seth and Giorgi, 1998) as it is this particular case. The decision also needs to take into account the available computational resources since increasing the domain size logically increases the computational costs (Leduc and Laprise, 2009). The compromise was solved by selecting a size of 2000 km x 2000 km centered on the CFN. For the simulation described herein, the selected area is displayed in Figure 3.1. A detailed evaluation of the influence of the domain size and location on the simulated wind field would certainly be an interesting topic for future research.

Once the size of the outermost domain was set, an appropriate horizontal resolution over the target area must be selected. The most suitable resolution is largely determined by the surface heterogeneity of the area under study and the specific purposes of the simulation. The CFN presents a relatively complex terrain (Fig. 1.2). Since the orography exerts a strong influence on the surface circulations, it would be desirable to use a high spatial resolution to represent it as realistically as possible. The reliability of the physical options available in the version of the WRF model used is uncertain as one approaches to horizontal resolutions of 1 km. For instance, as the horizontal resolution approaches the vertical depth of the PBL, the largest thermals are resolved to some extent, and the PBL schemes can introduce an over mixing (Stensrud, 2007). The radiative packages can also be problematic as the horizontal resolution approaches 1 km, since the available schemes represent the radiative processes along a single vertical column and the radiation passes through multiple horizontal grid cells as one approaches to resolutions of 1 km (Stensrud, 2007). A 3D radiative transfer scheme would be more appropriate in these situations (Cahalan et al., 2005). Thus, a 1 km grid resolution can be considered as a

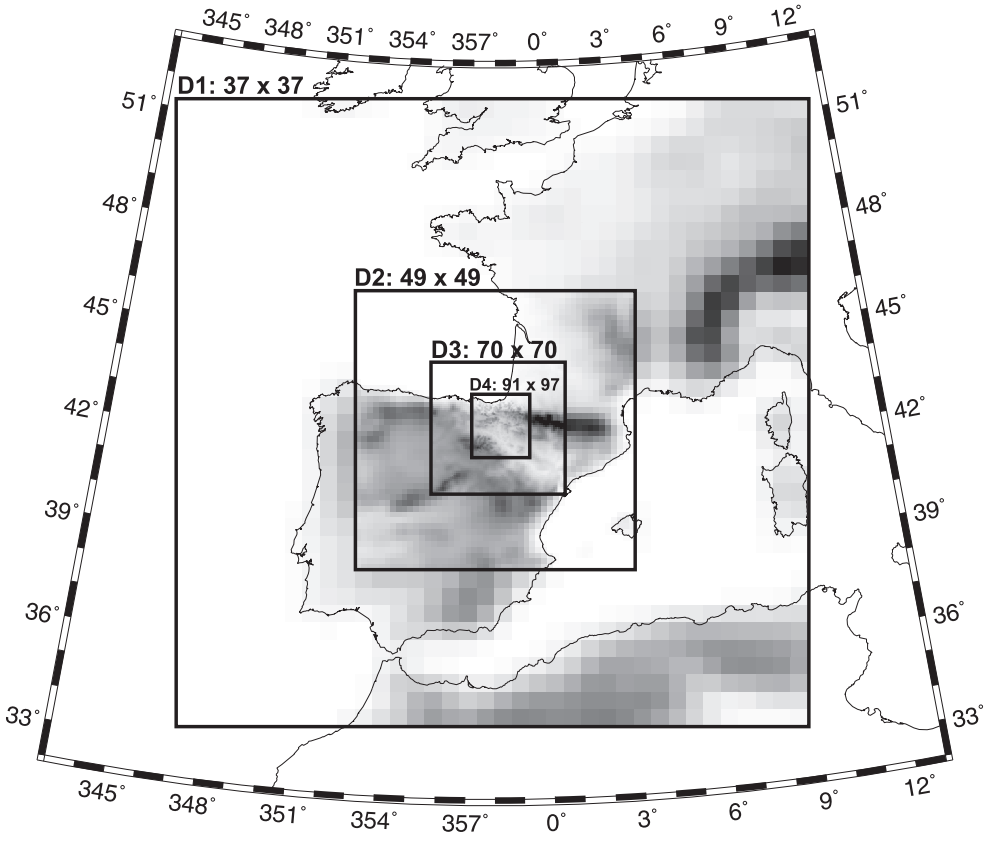

Fig. 3.1: Spatial configuration of domains used for the numerical simulation: four domains two way nested with 54 km, 18 km, 6 km and 2 km of horizontal resolution. The orography of each domain is displayed with their specific resolution. The number of grid points of each domain (West-East x Noth-South) is also displayed.

limit in the horizontal resolution achieved. Another factor that changes with resolution is the computational demand. Doubling the horizontal resolution requires four times more atmospheric volumes to simulate the atmospheric evolution over the same area; and the time step must be reduced by a factor of two in order to ensure stability on the calculations (Wicker and Skamarock, 2002; Skamarock et al., 2005). Thus, a simple calculation reveals that doubling the spatial horizontal resolution requires eight times more computational time. In the present experiment, a horizontal resolution of 2 km was finally selected as a balance between computational demand and need of physical detail on the simulation of the atmospheric evolution over the CFN. In comparison with a simulation at 1 km, this decision substantially reduces the computational time (by a factor of eight) without loosing too much resolution. Additionally, this resolution, while

avoiding to simulate under the always uncertain limit of reliability of a model, is still challenging in terms of the realism that the model can achieve.

The model is configured with four domains using two way nesting to reach the desired resolution over the CFN (Fig. 3.1). The outermost domain is centered on the CFN and has the size of 2000 km lat x lon and a horizontal resolution of 54 km. The rest of the domains are nested to progressively reach the desired horizontal resolution of 2 km in the innermost domain, which covers the whole CFN. The horizontal resolution as well as the number of grid points of each domain are displayed in Table **??**. The topographic data for all the domains is obtained from the USGS GTOPO30 dataset which has a grid spacing of 30" lat x lon (Bliss and Olsen, 1996; Gesch and Larson, 1996; Verdin and Greenlee, 1996).

A total of 31 levels were used in the vertical direction for all the domains. The top of the model is fixed at 50 hPa and the η levels at 0.0, 0.021, 0.046, 0.075, 0.108, 0.145, 0.188, 0.236, 0.290, 0.345,0.398, 0.451, 0.501, 0.550, 0.596, 0.639, 0.680, 0.718, 0.753, 0.786, 0.816, 0.844, 0.869, 0.892, 0.913, 0.933, 0.950, 0.966, 0.980, 0.993 and 1.000. The distibution of the levels in a vertical cross section is shown in Figure 3.2. These vertical levels are the default configuration of the WRF model and represent a commitment between the desirable high resolution at high levels to capture the large scale circulation dynamics, and the desirable high resolution at low levels to provide a better representation of the atmospheric boundary layer. For this study that focuses on the surface wind, it could be argued that it would be beneficial to add more levels within the PBL in order to provide a better representation of the vertical sub-grid scale fluxes. However, it has been reported that an acceptable representation of the PBL evolution can be achieved with a reduced number of vertical levels (Hong et al., 2006) and thus, it was decided not to increase the number of vertical levels since it would also increase the computational cost. In any case, a more in depth evaluation of the influence on the surface wind estimations due to an increment of the number of vertical levels close to the surface would be an interesting possibility for future research. Particularly interesting would be to analyze the influence of the location of the lowest level, since it defines the surface layer that is used to calculate the surface fluxes into the atmosphere.

3.3.2 Physical configuration

After setting a configuration for the domains, the physical schemes must be selected. The WRF model contains several options for each of the physical processes parameterized. The number of combinations of physical packages is therefore rather high. Some studies evaluated the sensitivity of a simulation to a particular parameterization (e.g. Zamora et al., 2003; Zhang and Zheng, 2004; Berg and Zhong, 2005). Other sensitivity studies evaluated the influence of different combinations of physical packages looking for the optimal configuration of the model (e.g. Fernández et al., 2007; Miao et al., 2008). These kind of studies reveals that the best set of schemes depends on the time of the day or the year, synoptic situation, simulated variable, area under study, etc. Some sets of packages outperform others under particular situations, but when the purpose is to simulate a long period that comprises a large number of atmospheric situations, as in the present

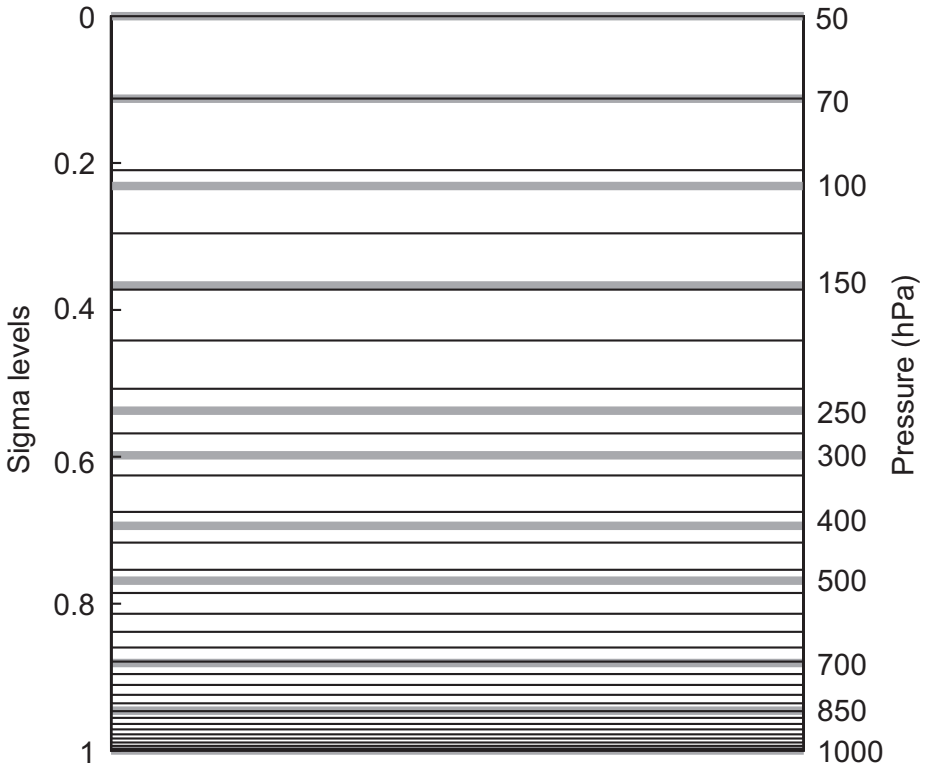

Fig. 3.2: Vertical levels used for the dynamical downscaling (black lines). The vertical levels from the ECMWF reanalysis/analysis data used to provide initial and boundary conditions are also displayed (gray lines).

study, the optimal set of physical schemes that systematically outperform the rest does not seem to necessarily exist.

On the basis of the previous considerations, the parameterizations schemes were selected based on experience and *a priori* considerations. The longwave radiation is represented with a Rapid Radiative Transfer Model (Mlawer et al., 1997) whereas the shortwave radiation scheme is based on Dudhia (1989). A modified version of the Kain and Fritsch (1990, 1993) scheme is used for the cumulus parameterization in the three outermost domains. The YSU PBL parameterization (Hong et al., 2006) is used in the four domains. With respect to the microphysics, the WRF Single-Moment 6-class scheme which is similar to (Lin et al., 1983) is adopted. Finally, a simple land surface model based on the MM5 5-layer soil temperature model is employed (Dudhia, 1996; Dudhia et al., 2004). The USGS land use/land cover system with a horizontal resolution of 1 km is used to determine the surface physical properties (Anderson et al., 1976). For more specific details on the parameterizations the reader is referred to Skamarock et al. (2005).

A further evaluation of the sensitivity of the surface wind simulation to any parameterization will remain as a future issue of research. This is not a trivial task since the performance of a parameterization is influenced by that of other parameterizations. For instance, the radiation scheme strongly depends on the cloudiness which is the target of the mycrophysics scheme. A very careful design of the experiment would be therefore mandatory to reach solid conclusions and, of course, these could vary for different meteorological situations. The schemes that should be the most important candidates subjected to evaluation, because of their prominent influence on the surface wind simulation, are the radiation, the soil model and the PBL scheme.

Finally, it should be keep in mind that the horizontal diffusion scheme selected operates over the η-coordinate surfaces that follow the terrain. This methodology can be problematic over the steep slopes of complex terrain areas as shown by the work of Zängl (2002) and Juang et al. (2005). The version of the WRF model used herein provides an option to perform the horizontal diffusion along geometric height coordinates but when it was adopted the simulation turned unstable. Updated versions of the model should be used on the future to evaluate the sensitivity of the surface wind field to this option.

3.4 Downscaling strategy

Initial and boundary conditions were obtained from the ERA-40 reanalysis project (Simmons and Gibson, 2000; Uppala et al., 2005) from 1992 to August 2002. After this date, up to 7 October 2005, the operational analysis at the ECMWF are used. The horizontal resolution of both datasets is 1^o lat x lon. The upper air information consist on geopotential heigh, temperature, relative humidity and the zonal and meridional wind components at the 11 vertical levles shown in Figure 3.2. At the surface, the information consist on 2 metre temperature and 2 metre dew point temperature, sea level pressure, 10 metre zonal and meridional wind components, and the soil temperature.

The WRF model is initialized as a "cold start" at 0 hours of each day, and is run for 48 hours, updating the boundary conditions with ECMWF data every six hours. The WRF simulation is stored every hour. The first 24 hours are discarded as model spinup retaining the outputs for the following 24 hours. The process was repeated until a simulation for the whole observational period was obtained (1 January 1992 to 7 October 2005). This sequence of short runs with numerous re-initializations as part of a longer simulation has been shown to outperform long-term continuous simulations with only one initialization (Pan et al., 1999; Qian et al., 2003; Lo et al., 2008) and is becoming increasingly accepted and adopted in recent years (Conil and Hall, 2006; Zagar et al., 2006). It can be argued, however, that it presents also some disadvantages. For instance, a longer simulation allows the model to equilibrate with its surface components providing a more interanlly consistent climatology (Giorgi, 2006).

The simulations were run in a cluster of PCs at the CIEMAT computing center. A single realization, to obtain a one day simulation as explained above, approximately elapses 40 hours of CPU. This means that more than 23 years would be necessary to run the fourteen years from 1992 to 2005 on a single processor. The subdivision of the simulation into short runs allows for a paralelization of the process and the complete

simulation was finished in about one year, (wall clock time). Presently, the simulation has been extended back to cover the period from 1960 to 2005.

4

Wind variability analysis: identification of areas of homogeneous wind behavior*

The present chapter is devoted to understand the surface wind variability over the CFN. The analysis is performed using the observational dataset presented in Chapter 2. There is a two fold objective for this analysis. First, to understand the wind behavior over the region. And second, to provide an appropriate framework to evaluate the capability of the dynamical downscaling presented in Chapter 3 to reproduce the observed wind variability. The latter will be the focus of the the next chapter.

The surface wind variability is analyzed by dividing the CFN into a small number of subregions with internally homogeneous wind behavior, by means of a wind regionalization. This allows to inspect the spatial wind variability by analyzing the distribution of the subregions identified, and the temporal wind variability by analyzing the wind behavior at the different subregions.

4.1 Regionalization background and necessity of a reduced dataset

The most popular strategies to identify areas with homogeneous behavior are based on multivariate analysis (Chatfield and Collins, 1980). These techniques have been successfully applied in a variety of case studies and for several variables (e.g. Dyer, 1975; White et al., 1991; Coronato and Bisigato, 1998; Bonell and Sumner, 1992; Fovell and Fovell, 1993; Sanchez-Lorenzo et al., 2007, 2008; Burlando et al., 2008). However, the major efforts have been oriented to identify climate zones (e.g. Gadgil and Joshi, 1983; Stooksbury and Michaels, 1991; Fovell and Fovell, 1993; DeGaetano, 1996; Bunkers and Miller Jr., 1996; Malmgren and Winter, 1999) or affinity areas of precipitation (e.g. Dyer, 1975; Bärring, 1988; Ogallo, 1989; Mallants and Feyen, 1990; Bonell and Sumner, 1992;

* The main contents of this chapter are included in:
Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, J. Navarro, E. García-Bustamante and F. Valero, 2008: Surface wind regionalization in complex terrain. *J. Appl. Meteor. & Climatol.*, **47**, 308-325.

Fernández-Mills, 1994; Gong and Richman, 1995; Comrie and Glenn, 1998; Romero et al., 1999b; Muñoz-Díaz and Rodrigo, 2004).

The potential of the regionalization approach has seldom been explored with wind related variables (Kaufmann and Weber, 1998; Cheng, 1998; Burlando et al., 2008). Kaufmann and Weber (1998) applied cluster analysis (CA) to the correlation matrix of the wind direction finding wind regions related to the orographic features of the area under study. Cheng (1998) identified homogeneous regions of maximum wind speed using basic probability functions, homogeneity tests and composite frequency functions. Burlando et al. (2008) applied several CA approaches to wind speed observations acquired over Corsica to compare the different classifications and thus find the most consistent wind speed regions. None of these previous studies consider the vectorial nature of the wind, since they classify scalar fields such as wind direction, maximum speed or wind speed.

The regionalization of the surface wind over the CFN is herein performed using multivariate techniques in order to exploit its potential to manage a vectorial variable such as the wind. Two different methods based on principal component analysis (PCA) are used in this study to regionalize. The first method carries out CA of the most important PCA modes (Romero et al., 1999c) which allows identifying homogeneous wind climate variability groups, while the second method makes use of rotation of selected principal components (White et al., 1991). These are standard methodologies used to regionalize, but they have not been previously applied to the horizontal wind vector. The use of two different methodologies allows for the comparison of results from both procedures in order to reach the most consistent classification. The wind regionalization analysis is performed with a previous version of the observational dataset described in Chapter 2. The observations were taken at the same locations but this previous dataset spans the period from 1 January 1992 to 30 September 2002.

The observational time series present cosiderably high number of missing records in the time series mostly associated with a later instalation of the sensors (Fig. 2.21). The existence of missing data and thus of uneven time intervals can potentially produce adverse effects in the calculation of principal components (Yarnal, 1993). This can be partially mitigated by interpolating the data into a grid (Wilks, 1995). However, grid interpolation of wind data can be particularly problematic in complex terrain regions where exposure, orographic features, and altitude are usually more important than distance (Kaufmann and Weber, 1998; Steinacker et al., 2006). In addition, interpolation does not incorporate new information to the dataset unless more variables or sites are considered. An alternative possibility is to undertake a careful *pairwise* treatment of missing values as suggested by Bärring (1988). Ludwig et al. (2004) applied PCA both to grid interpolated and unevenly spaced station data for the same region obtaining similar results; other examples of PCA with unevenly spaced data are those of Bonell and Sumner (1992) and Comrie and Glenn (1998).

This last approach is here adopted by applying PCA to a subset of data with fewer missing data in its daily fields in order to secure a more robust estimation of eigenvectors. The complete dataset is subsequently used to calculate an extended version of the principal components which allows for studying the variability in each region at longer time scales (see Section 4.4). The subset of data is created by selecting only the 35 stations with the best quality wind measurements, in terms of the availability of data, from the

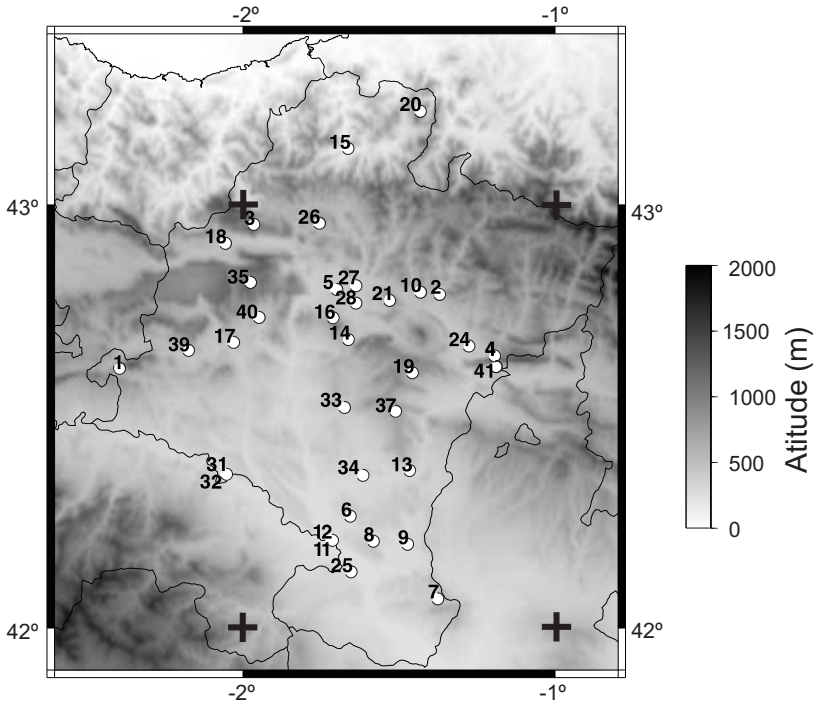

Fig. 4.1: Location of wind stations within the CFN. Shading represents altitude, circles are the measurement sites and the thin lines highlight political boundaries. See Table 2.1 for specific station descriptions. The crosses represent the grid points within the ECMWF reanalysis dataset.

meteorological network of the CFN (Fig. 4.1). In addition, only the daily fields with more than 80% of the site observations available (i.e. more than 28 stations) are retained in the reduced datset. The number of available daily fields is in this way reduced to 947, but the reduced dataset ensures a homogeneous representation of all stations during the time steps for the calculation of eigenvectors. A further selection of fields was additionally made considering the resulting monthly distribution of available daily fields (Fig. 4.2). The irregular spread of data over the year could potentially stress the prevailing circulations of the months with more available daily fields. This potential undesirable effect was weakened by imposing an upper limit on the retained number of daily fields for each month, thus achieving a more homogeneous distribution. This threshold was established in this case selecting the best quality 65 daily cases for each month (dashed line in Fig. 4.2), the only exception being February which could only accumulate 54 daily fields. The final subset containing a total of 769 daily wind measurement fields was employed for the wind regionalization. Since the quality of the original dataset was progressively improved with time (Fig. 2.21b), the latter subset was mostly concentrated in the recent years spanning the period from October 1999 to September 2002.

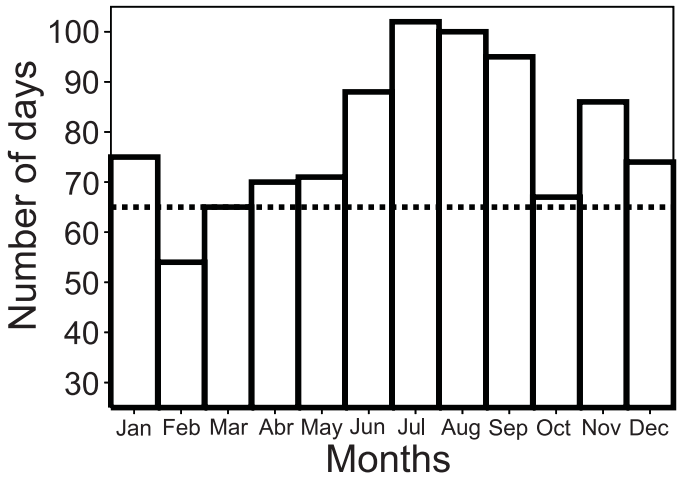

Fig. 4.2: Monthly distribution of the number of days with more than 80 % of wind observations available (boxes). The dashed line represents the final homogeneous distribution after discarding the days with larger amounts of missing values from each month (except for February that has only 54 available days).

4.2 Regionalization methodologies

The two selected regionalization methodologies perform a PCA in a first step. Due to the vectorial nature of the wind, a scalar PCA can be applied to its components (Barnett, 1977) or a vectorial PCA to the wind vector. The latter analysis can be faced from two different approaches: a real vectorial PCA (Ludwig et al., 2004; Jiménez et al., 2008) or a complex vectorial PCA (Hardy, 1977; Hardy and Walton, 1978; Horel, 1984). In comparison to the scalar PCA the vectorial PCA presents the advantage that the relationship between both components is not necessarily lost during the analysis (Klink and Willmott, 1989). Furthermore, the real vectorial PCA has been reported to be more effective than the complex one in discriminating the flows corresponding to different temporal/spatial scales that are orthogonal (Kaihatu et al., 1998). For these reasons, the real vectorial PCA was selected for this study, and applied to the correlation matrix of available zonal and meridional time series (S-mode; see Richman 1986) as a first step in both regionalization methods. The use of the correlation matrix, instead of the covariance matrix allows for the comparison of sites with different ranges of wind variability. PCA is a standard methodology and the reader is referred to text books for more information (Preisendorfer, 1988; von Storch and Zwiers, 1999). A brief definition is presented in Eq. (4.1):

$$(u(x,t),v(x,t)) = (\bar{u}(x),\bar{v}(x)) + \sum_{i=1}^{2N} a_i(t)(u_i(x),v_i(x)) \tag{4.1}$$

where the elements on the left hand side are the zonal u(x,t) and meridional v(x,t) wind data at time t and at each of the N sites (x=1,2,...,N). $\bar{u}$ and $\bar{v}$ represent the time average of the zonal and meridional wind components. The second term on the right hand side involves the $i = 1 \ldots 2N$ principal modes through the product of the time series of scores ($a_i(t)$, principal components) and the eigenvector ($u_i(x), v_i(x)$). The selection of the number of principal modes to retain, M, was made with the classical scree test of Cattell (1966).

4.2.1 Cluster analysis methodology

A CA of the retained PCA modes allows for obtaining groups of stations with similar loads. The target of any CA is to form groups with large similarities between objects inside a cluster and small similarities between objects of different clusters. Hence, the first step to perform a CA is to define a metric for similarity between the objects to be grouped. In this analysis the objects are the stations and the measure of similarity between two given stations A and B is defined by:

$$d_{AB} = \frac{1}{M} \sum_{i=1}^{M} ([u_i(x_A) - u_i(x_B)]^2 + [v_i(x_A) - v_i(x_B)]^2)^{\frac{1}{2}} \tag{4.2}$$

where i denotes the principal mode and u and v are the components of the vectorial loads as in Eq. (4.1). Once the similarity measure has been defined, a clustering procedure must be chosen. Hierarchical and non-hierarchical algorithms stand as standard tools for this purpose (Anderberg, 1973). Hierarchical techniques are divided into divisive and agglomerative; divisive procedures start with all the objects in a cluster and split them at each iteration of the procedure while agglomerative procedures consider initially each object as a cluster and merge the two most similar elements at each step. Divisive procedures are slower than agglomerative (Fovell and Fovell, 1993) and this is probably why they are less popular. All the hierarchical procedures present the same disadvantage: if an object is erroneously assigned to a group at the beginning of the clustering process, this object cannot be relocated (Anderberg, 1973). The nonhierarchical procedures outperform the hierarchical schemes (Gong and Richman, 1995), however, a drawback is that they need to know in advance the number of clusters to be formed and their initial seed values. Hence, a two step cluster procedure (Milligan, 1980) is frequently used to solve these problems (Stooksbury and Michaels, 1991; Kaufmann and Weber, 1996; Kaufmann and Whiteman, 1999) and was the method selected in this work. In the first step a hierarchical algorithm finds the groups which serve as an initial state for a second phase in which a non-hierarchical scheme reorders the objects. Weber and Kaufmann (1995) compared different hierarchical methods for a wind field classification adopting the distance defined by Eq. 4.2. They found some advantages in the Complete Linkage Algorithm (CLA Johnson, 1967) over the rest of the methods tested. Thus, this was the clustering procedure selected for the first step. This algorithm defines the degree of similarity between two clusters as the largest value of the distance in Eq. (4.2) between any possible pair of objects, each one extracted from each cluster. With this definition, CLA merges two clusters only if the distance between the most dissimilar objects does

not exceed a certain threshold. In this way, the method ensures broad similarity of the elements within a cluster and dissimilarity of these against those of other clusters. In order to decide on the appropriate number of clusters to form, the value of the similarity measure at which the two most similar clusters are merged at each step is screened; a large change in the value of the distance before and after merging two potentially similar groups means that two very different clusters have been merged and can be used as an indication to stop the algorithm in the previous step.

On the second CA step, a method similar to the non-hierarchical k-means procedure (Kaufmann and Weber, 1996) has been used. This algorithm calculates the similarity based in Eq. (4.2) between each station and a reference centroid representative of the cluster to which the station has been previously assigned. An initial cluster assignment is made in the previous step with the CLA and centroids are calculated as the average of all individuals within each initial group. After this, distances according to Eq. (4.2) are calculated on the basis of loadings of each target station and those of the centroid. This allows for a new reassignment through which stations can be moved to a different group that presents minimum distance between its centroid and the target station. Once the procedure has been applied to each site, new reassignment steps can be undertaken in an iterative manner until stability is attained and no station is virtually relocated in a different group.

4.2.2 Rotation of principal components methodology

The second regionalization approach explored herein is based on the rotation of the selected principal modes to obtain the wind regions. The aim of the rotation is to produce 'simple structure' in which the variables are as close as possible to a hyper-plane of at least one principal mode (Richman, 1986; Hannachi et al., 2006). With the simple structure, the loading map of each principal mode weighs on a different subregion.

Theoretically, each observational site should present a high load in just one rotated loading map and null loads in the rest (perfect simple structure). This would lead to a situation which each map would define a completely different subregion. Actually, the loads are not null and in practice it is necessary to define a *critical* threshold value to define the subregions: only those sites with loads higher than the critical value will belong to an specific subregion. Since a vectorial PCA is adopted the loads are vectors and the critical value is defined based on the value of their module.

Rotations can be either orthogonal or oblique. There is considerable discussion on the benefits and disadvantages of each type of rotation. Some authors find very similar groups with both methods (Gregory, 1975) while others conclude that the oblique rotations produce more stable results and are superior to the orthogonal rotations (White et al., 1991). VARIMAX (orthogonal, Kaiser, 1958) and OBLIMIN (oblique, Clarkson and Jennrich, 1988) rotation techniques were tested and little difference was found for the case of this study. The first technique was finally selected because of its simplicity and the property of preserving the orthogonality of the eigenvectors after the rotation.

This regionalization method allows for one station to belong to more than one subregion since the regions derived from each PCA mode can overlap, contrary to the CA

method that generates a hard regionalization in which each station belongs to only one subregion.

4.3 Identification of the wind subregions

As a preliminary inspection of the wind variability over the region, the average and standard deviation fields of wind speed module are displayed in Figure 4.3. The spatial patterns of both variables suggest a linear relation with the sites showing higher wind averages also presenting more variability. This is typical of positive defined variables like precipitation (Xoplaki et al., 2004) in which increases in mean values lead to wider probability distributions (higher variability). For the case under consideration Figure 4.4a illustrates this relationship (correlation r = 0.98). The sites with the strongest wind are the mountain stations and some of the stations in the Ebro valley (Fig. 4.3). Many of the windy sites are also the highest in altitude (Table 2.1) suggesting a relation between altitude and wind speed. This can be better observed through the dispersion diagram of both variables (Fig. 4.4b) that reveals a correlation value of 0.78.

Mean wind vectors are displayed in Figure 4.3 calculated from averages of the zonal and meridional wind components. The mean flow is from the northwest and is channeled from the northern valleys to the Ebro valley through the North-South passages around the eastern mountains. The meridional component presents, in general, higher variability than the zonal component (Fig. 4.4c). The correlation between both components (r = 0.76) indicates some linear relationship between them. This can be argued to be a reasonable feature from the point of view of the conservation of horizontal momentum. Assuming the loss of energy from surface friction is small and loss of horizontal momentum due to vertical ascents or descents represent a small fraction of horizontal momentum on daily time scales, the changes in the zonal (meridional) component due to interactions of the wind with orographic obstacles will translate into transfer of momentum to the meridional (zonal) component. Thus, changes in the zonal (meridional) wind component will often be related to changes in the meridional (zonal) component and ultimately sites showing higher variability in one component should be expected to have also higher variability in the other component. This relation is an indication of common variability and supports the joint treatment of both components in the application of the vectorial PCA in the analysis of the wind variability instead of performing a separate analysis on each variable component. As for the higher meridional than zonal variability, this is related (not shown) to the channeling of the flow between the large mountain systems in northern Spain (the Cantabrian mountains and the Pyrenees, Fig. 1.2) and at a more regional scale within the CFN. Here, channeling is favored along the northern valleys and the Ebro Valley with a NW-SE orientation and particularly around the eastern mountains systems with a N-S orientation. This behavior will be further illustrated with the results of the PCA analysis in the following sections.

4.3.1 Regionalization using cluster analysis

The explained variance of the leading PCA modes of the wind field is shown in Figure 4.5. There are breaks of the slope at mode three, five as well as a less clear one at

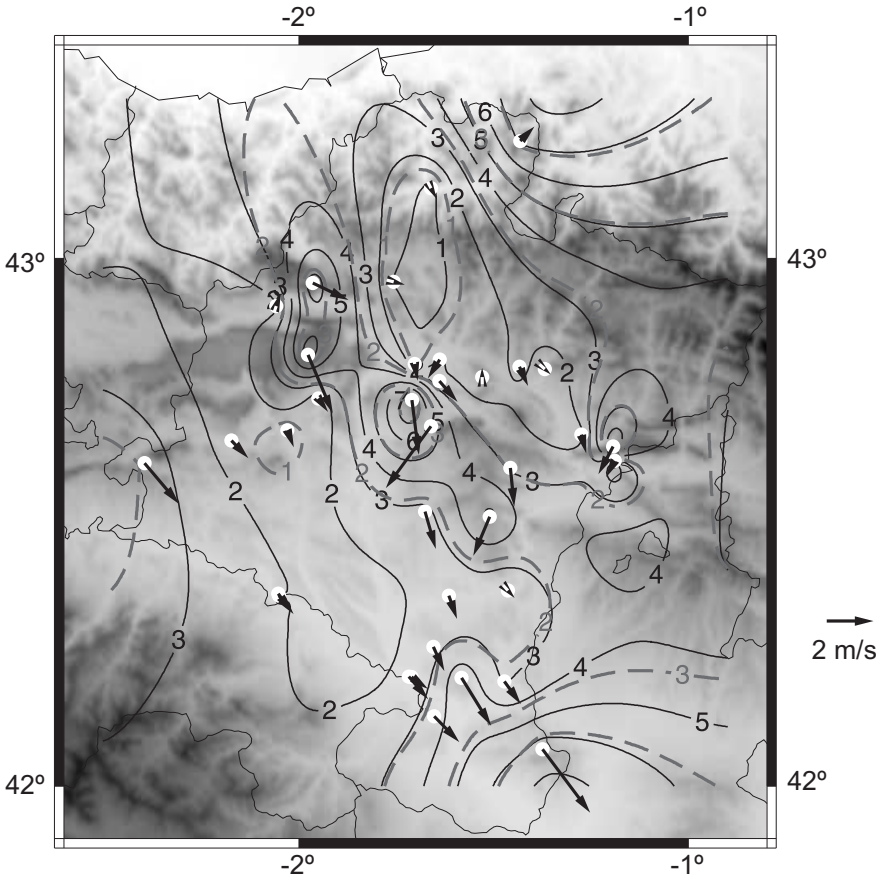

Fig. 4.3: Mean wind speed field (solid lines), its standard deviation (dashed lines) and mean zonal and meridional wind components (vectors) of the selected daily wind fields. The length of the vectors is therefore proportional to the module of the mean wind vector. The tail of vectors is placed at the observational sites.

mode nine which according to the scree test of Cattell (1966) are reasonable numbers of modes to retain. The retention of three of them might not be enough to adequately group the stations because in such cases the similarity measure (4.2) would be calculated with the contributions of only three terms whereas, nine modes could introduce noise into the classification because of the little variance explained by the higher order modes. Therefore, five principal modes which accumulate 84.1 % of the variance in the data were retained. The loading maps (eigenvectors) of these principal modes are displayed in Figure 4.6. These maps represent flow directions and can be interpreted in their positive phase as displayed in Figure 4.6 or in the opposite sign of the mode (corresponding to

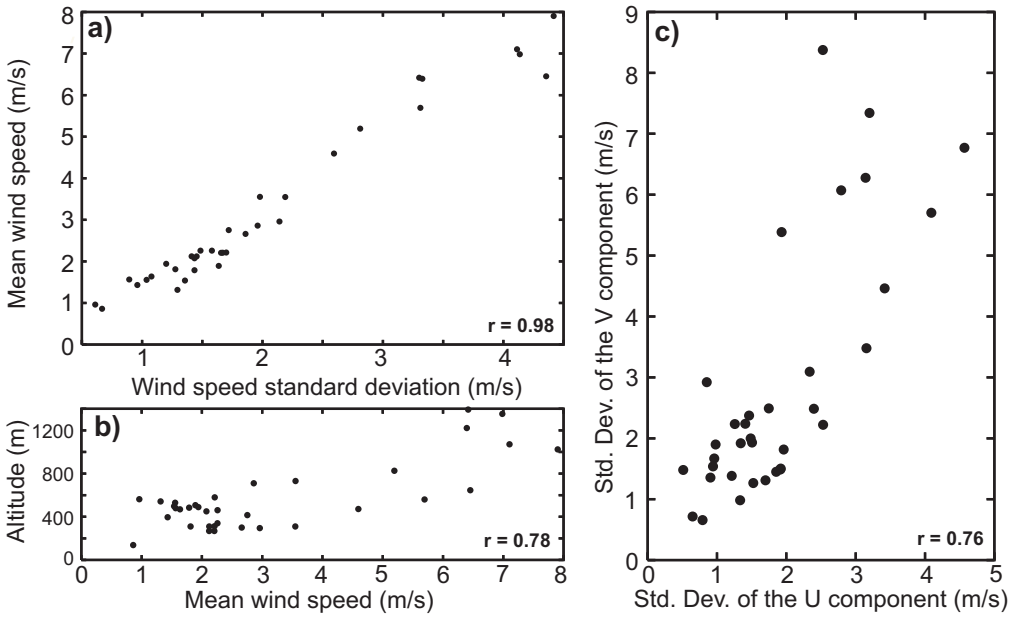

Fig. 4.4: a) Mean daily wind speed *vs.* its standard deviation for the selected days in the 35 sites, b) altitude *vs.* mean wind speed and c) standard deviation of the v wind component *vs.* standard deviation of the u wind component.

a negative sign in their principal component). The first mode explains two thirds of the variance (66.8 %) and is well organized (Fig. 4.6a): the vectors are aligned along the valley axis indicating that the dominant physical process is the topographical channeling. In the positive (negative) phase the main flow has a SE to NW (NW to SE) direction while some mountain sites seem to show a certain decoupling of the flow with respect to the valley circulations presenting a meridional direction with a northward (southward) orientation. The second principal mode explains 8.6 % of the variance and presents strong eastward (westward) flows at the highest locations mainly, and weaker SW to NE (NE to SW) flows at the rest of the sites (Fig. 4.6b). Physically it can be interpreted as the influence of the synoptic scale flows which is more intense at the higher sites. Therefore, this mode also reveals a decoupling of the flow between mountain and valley circulations noticed before. The third principal mode explains 3.4 % of the variance. It shows activity in the center of the region displaying zonal flow directions but with opposite orientations at sites separated by relatively small distances that could be associated with re-circulations or local behaviors of the flow (Fig. 4.6c). The fourth and fifth modes explain 2.9 % and 2.4 % of the variance respectively and show relations between a few stations of each mode (Fig. 4.6d and e). The low percentage of variance accounted for these last modes can be related to the limited size of their area of influence relative to that of the three first modes. Improved knowledge of variability within this confined areas could stem in the

future from higher spatial sampling. This arrangement of explained variance is similar to that found by other authors in regions of comparable size (Hardy and Walton, 1978; Green et al., 1992a; Ludwig et al., 2004).

The first step of this CA regionalization is to group together the stations with similar loads (Fig. 4.6) using the hierarchical CLA in order to decide the number of subregions to form. This was done displaying the sequence of distance measures at which the clusters were merged in each step (Fig. 4.7). Since the CLA merges the two most similar clusters at each step, a large jump in the sequence means that two very different clusters have been merged and indicates the convenience of stopping the algorithm just before this happens. Steps appear at the case of six and nine clusters (Fig. 4.7) and therefore these seem reasonable number of subregions to form. The regionalization of nine subregions forms many small groups with only one or two stations (Fig. 4.8a) hence a six-cluster regionalization was selected (Fig. 4.8b). After the CLA, the reordering of the stations in the six selected subregions is performed by the non-hierarchical algorithm and provides the final wind regionalization (Fig. 4.8c). This reordering only changes the location of station number 37 (Fig. 4.1), which was assigned to the first subregion by the CLA (Fig. 4.8b) and now belongs to the third wind region (Fig. 4.8c). The similar clustering of the stations obtained in the two steps of the CA methodology grants robustness to the proposed regionalization. The Ebro valley stations form the first subregion (label 1 in Fig. 4.8c), the narrow northern valleys the second (label 2 in Fig. 4.8c), the high mountain stations the third (label 3 in Fig. 4.8c) and the rest of the subregions are small groups with a North-to-South orientation in the center of the CFN (labels 4, 5 and 6 in Fig. 4.8c). Conceptually, this seems to suggest that there

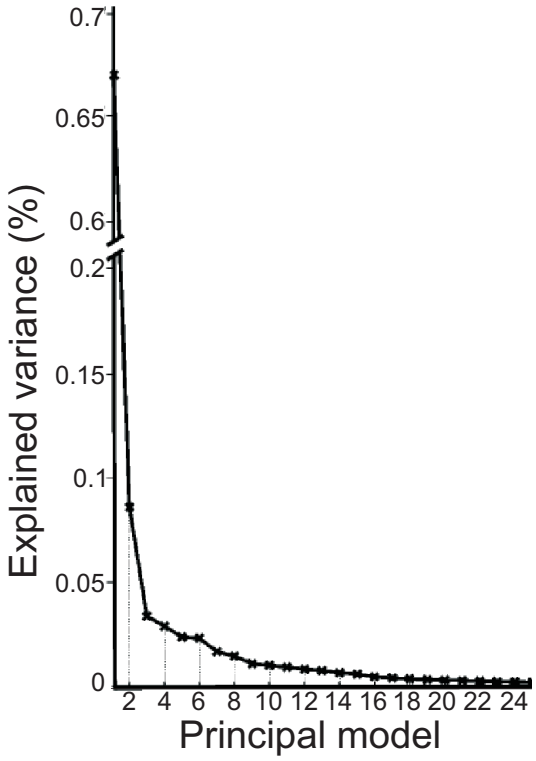

Fig. 4.5: Explained variance of the PCA modes. Notice the break in the vertical scale.

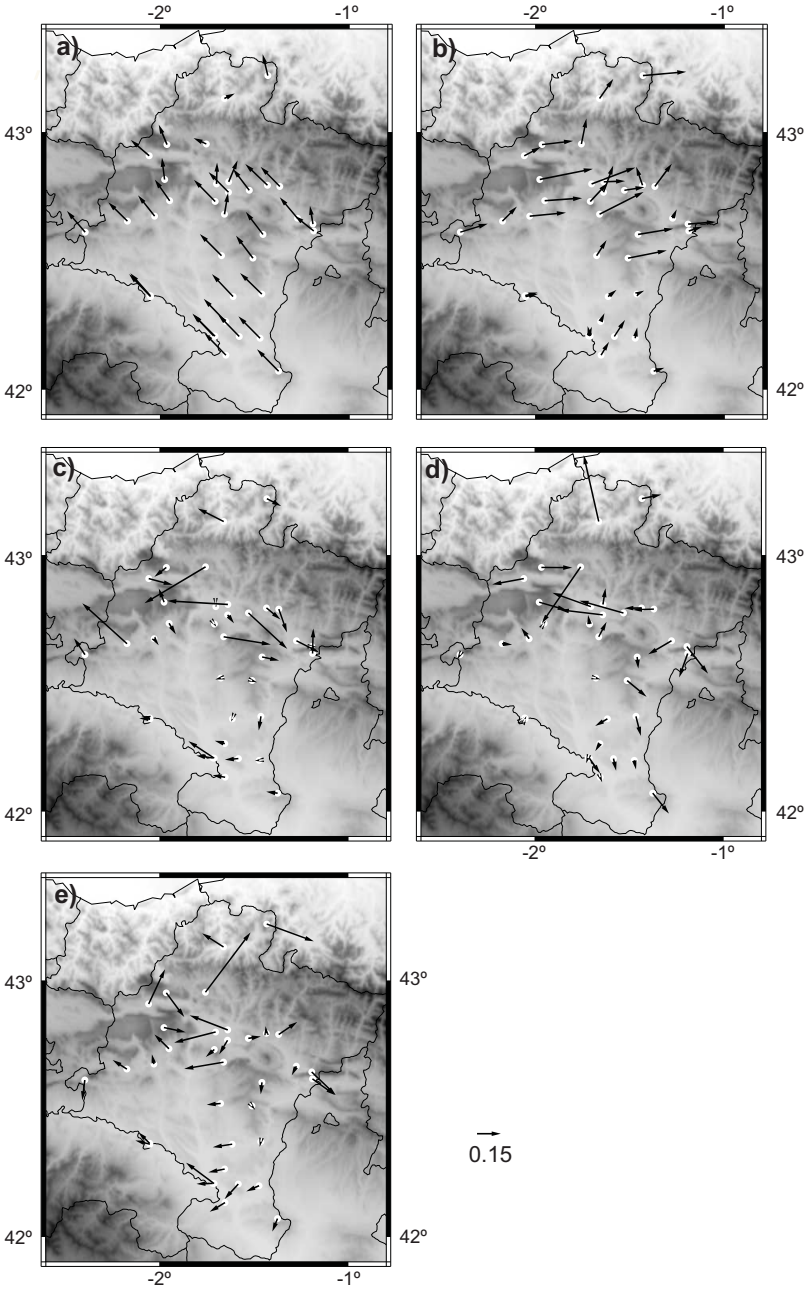

Fig. 4.6: Loading maps of the first five principal modes. The variance accounted for by each mode are a) 66.8 %, b) 8.6 %, c) 3.4 %, d) 2.9 % and e) 2.4 %.

are three well defined regions, and a fourth one, the central North-to-South area which groups the clusters with lowest numbers of stations.

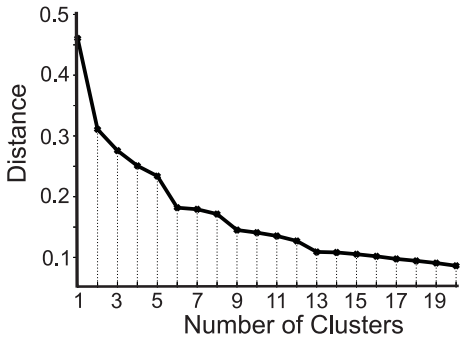

Fig. 4.7: Distance at which the last two clusters are merged against the number of clusters formed. The larger jump at 6 and 9 suggests to retain these numbers of clusters.

4.3.2 Regionalization using rotated principal modes

This method rotates the retained PCA modes to perform the regionalization. Subregions are then formed retaining just the highest loads of each loading map. This is done by defining a critical threshold value and, only those sites with loads exceeding it will belong to the subregion. Hence, one subregion is created for each rotated mode and the method allows for the formation of overlapping subregions in opposition to the CA method which generates a hard regionalization without possible overlap. The retention of five modes, as was done with the regionalization using CA, results in a strong overlapping between the subregions formed and suggests that five wind regions appear to be too many. Hence, only four principal modes were employed for this regionalization method (81.7 % of the variance). As was mentioned above, rotation tends to form *simple structures* (Richman, 1986) in which the variables are as close as possible to the hyper-plane of at least one principal mode. The degree of the simple structure can be visualized displaying the loads of one principal mode against those of another mode. Figure 4.9 shows an example of this for the first mode 1 *vs.* 2 and 4 illustrating that values overall tend to be closer to the axis in the rotated case. The effect off the VARIMAX rotation is apparent in the turning of the dispersion diagram of mode 1 *vs.* 4.

The loading maps of the four rotated principal modes in their positive phase are displayed in Figure 4.10. Compared with the unrotated principal modes (Fig. 4.6) they show in some aspects a clearer physical interpretation. The first and second rotated principal mode (Fig. 4.10a,b) present similar patterns to those of the first two modes (Fig. 4.6a,b). The first pattern presents in its positive (negative) phase the SE to NW (NW to SE) channeling along the Ebro Valley, as in the unrotated case, and some minor differences in the representation of its influence in the area between the western and eastern mountains,

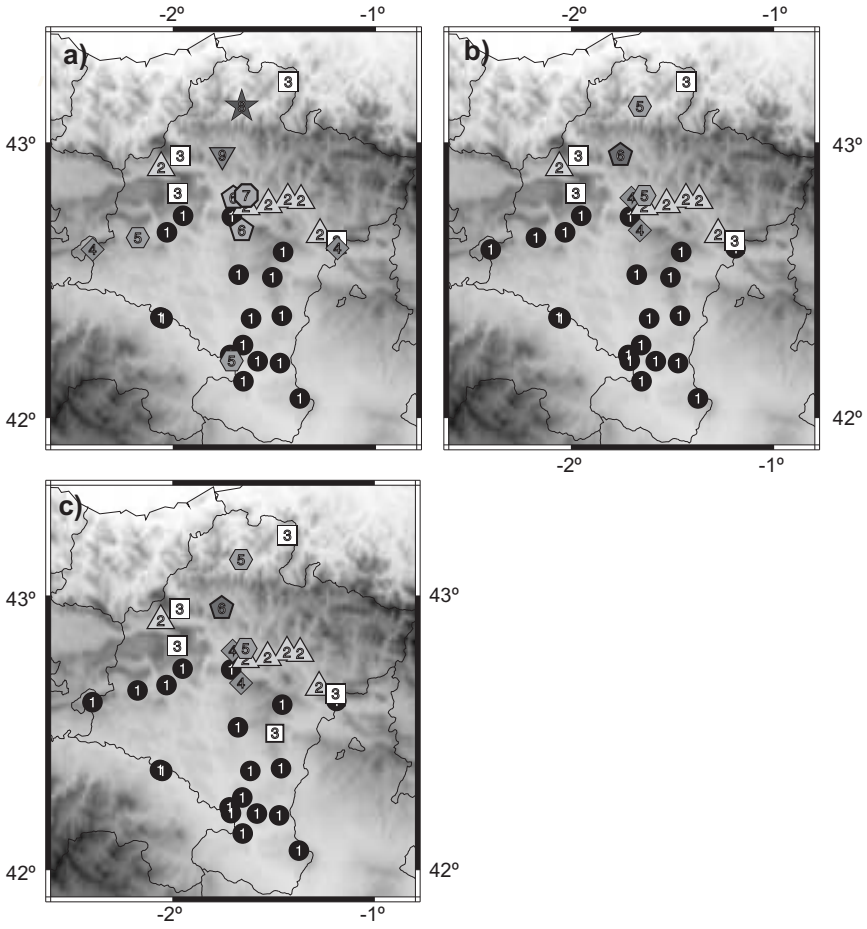

Fig. 4.8: Wind regionalization obtained with the CLA of the five most important principal modes for the cases of 9 (a) and 6 (b) subregions; c) shows the wind regionalization obtained with the second step of the CA methodology, the non-hierarchical algorithm that reorders the 6 subregions formed with the CLA (b).

where the rotated mode indicates perhaps some clearer tilt to the North (South). The second mode (Fig. 4.10b) is also similar to its unrotated analogue (Fig. 4.6b) but with the vectors presenting a clearer alignment along the eastward (westward) direction in the mountain stations and with weaker vectors (smaller loadings) in the valleys. The third rotated principal mode (Fig. 4.10c) presents meridional orientation of the vectors that show southward (northward) sense in the eastern areas of the CFN, and zonal orientation with westward (eastward) sense in the western areas. In this case the pattern suggests a clearer behavior of the wind flow bordering the western and eastern mountain *obstacles*

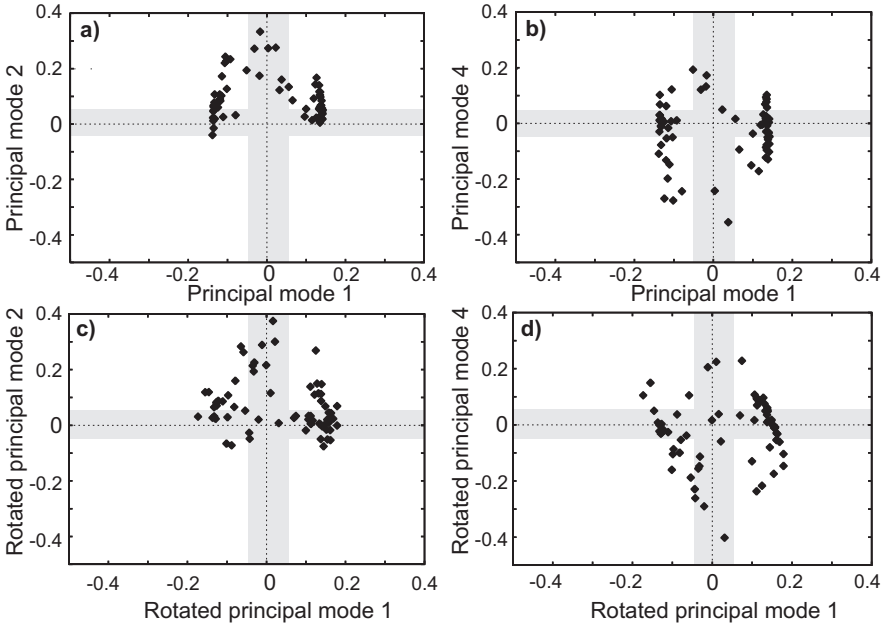

Fig. 4.9: Dispersion diagram of the rotated 1st principal mode plotted against modes 2 and 4.

from some North or NE direction. A further analysis of the synoptic conditions related to the surface circulations will be presented in Chapter 6 to understand the regional behavior as a result of interaction of large scaled dynamics with topography. The fourth rotated principal mode (Fig. 4.10d) shows channeling flows along the northern valleys and weak vectors in the Ebro valley and the mountain stations. In comparison to the last unrotated vectors (Fig. 4.6) it highlights distinctly the channeling of wind in the northern valleys.

After a visual inspection of the loading maps (Fig. 4.10) a critical value for the vector modules can be defined in such a way that the stations that exceed the threshold will define wind regions. As it can be observed in Figure 4.10, the critical value must be chosen carefully in order to avoid too much overlapping between subregions or stations left ungrouped. Several critical values were tested and finally, the compromise was solved by adopting a critical module value of 0.175 which delimits the most consistent subregions. The regionalization obtained can be observed in Figure 4.11a. The first subregion corresponds to the Ebro valley (EV, circles in Fig. 4.11a); the second is mainly defined by the mountain stations (MS, squares in Fig. 4.11a); the third region groups stations lined up in a North-South direction from the inner northern valleys up to the Bidasoa mountains and beyond (NS, diamonds in Fig. 4.11a) and the fourth is fundamentally shaped by the northern valley sites (NV, triangles in Fig. 4.11a). This wind regionalization is very similar to that obtained with the CA method (Fig. 4.6c). However, it clearly groups the sites labeled with diamonds, region 3, as a whole region overlapping with the northern

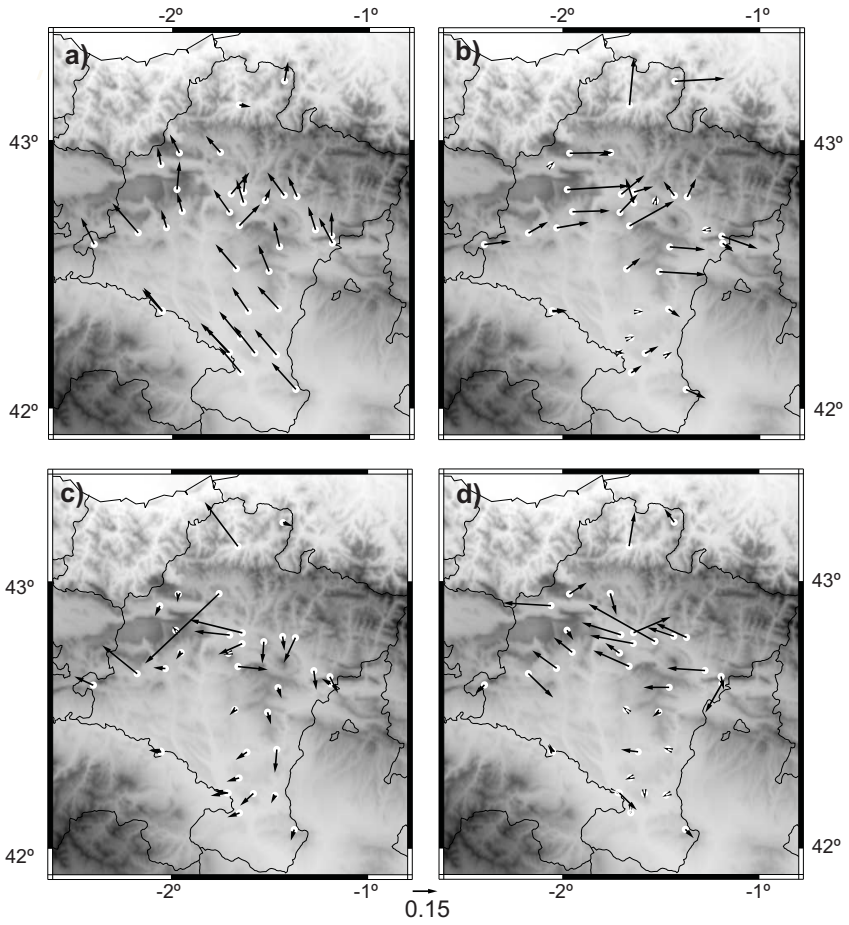

Fig. 4.10: Loading maps of the four rotated principal modes with the VARIMAX technique. The tail of the vectors are placed at the observational sites. See Table 4.1 and Fig. 4.12 for complementary information.

valley sites, instead of the various small groups provided by the CA. Moreover, this regionalization creates an extended group of mountain stations instead of the very specific individual high mountain sites identified by the CA regionalization method. However, this methodology has the drawback that there are two stations, 10 and 16 (Fig. 4.1), that were not assigned to any group. When the critical value was reduced in an attempt to include them into a group then the subregions overlapped too much. Station 16 has the highest load module (0.168) for the EV subregion, which agrees with the assignment resulting from the CA regionalization; however, it also has a high load for the MS region (0.155). Station 10 has the highest load modules for the EV subregion (0.171) and the

NV subregion (0.169). These type of situations could be ameliorated with further improvements in the temporal length and spatial coverage of the dataset which would allow for capturing better signals which are faintly represented in the data.

The two regionalizations present their own drawbacks and advantages and the use of both methodologies allows to reach more robust results. The grouping of stations provided by both methodologies (Figs. 4.6c and 4.11a) will be used to evaluate the ability of the simulation in reproducing the spatial wind variability in the next Chapter 5. The definition of subregions from the rotational approach (Fig. 4.11a) will be used to calculate the regional wind time series since it shows a simpler structure than the CA based approach (Figs. 4.6c), with no group formed by just one station. Hereafter, when referring to the regional time series we will allude to the averaged series of the stations belonging to each of the subregions isolated in the rotational approach (Fig. 4.11a). The rotational scheme also presents the advantage that makes it possible to examine the temporal wind variability of each subregion by analyzing the time series of the rotated scores. This will be analyzed in the following section.

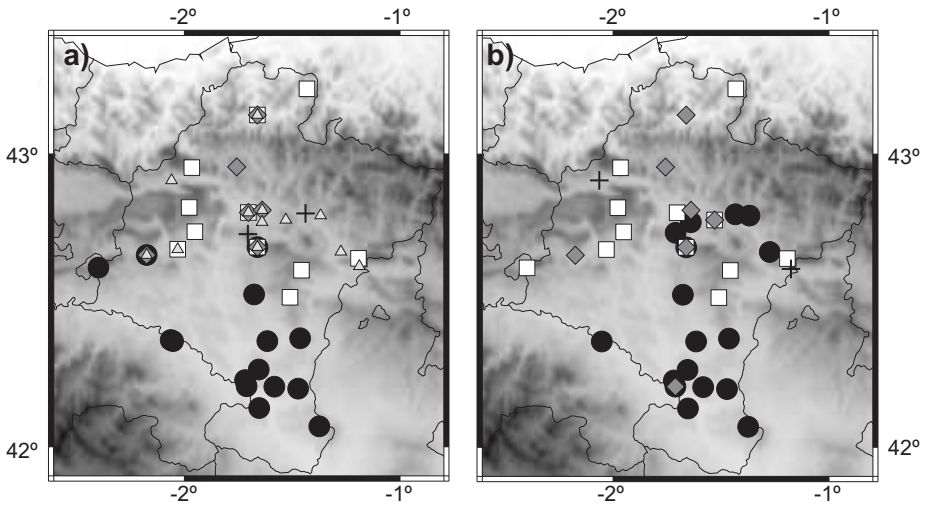

Fig. 4.11: Wind regionalization obtained with the rotation of the four (a) and three (b) most important principal component modes. The crosses represent the unclassified stations.

4.4 Temporal wind variability at the subregions

The 20 days moving average filter outputs of the time series of scores are displayed (Fig. 4.12). Each score presents different behavior showing the distinct wind variability in the subregions formed. However, the rotation of the principal modes involves the loss of the

independence (zero correlation) property of the scores time series and therefore, part of the variance accounted by one mode could also be explained by other modes (Preisendorfer, 1988). The correlation among the rotated scores time series can be observed in Table 4.1. The correlation of 0.79 between the score 1 (EV) and the score 4 (NV) implies similar wind variability in these subregions. Indeed, the regionalization reached retaining three principal modes (Fig. 4.11b) basically groups the stations from these two subregions. However, the fourth subregion in Figure 4.11a is a well defined cluster using the CA regionalization methodology (Fig. 4.8c) and is fundamentally constituted by stations that share a characteristic terrain feature, namely, to be located at the northern valleys. Thus, the decision was taken to keep this region as a different group on the basis of the results with the CA method and the distinct topographical character.

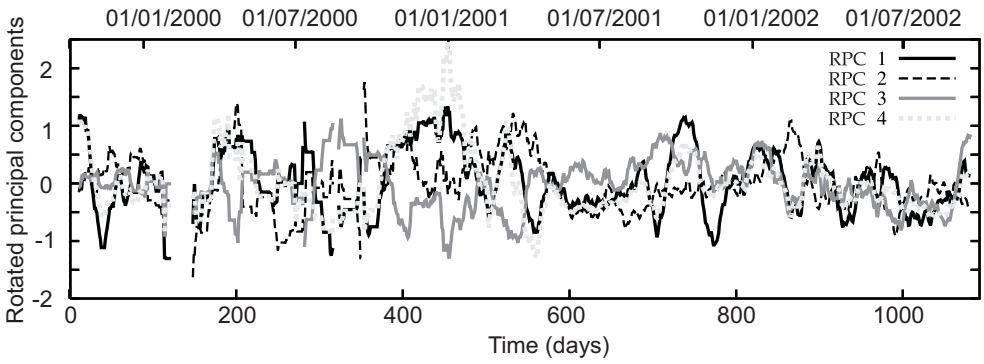

Fig. 4.12: 20 days moving average filter outputs of the time series of scores after VARIMAX rotation.

Table 4.1: Correlation of the rotated scores.

	Score 1	Score 2	Score3	Score 4
Score 1	1.00	-0.25	0.06	0.79
Score 2	-0.25	1.00	-0.29	-0.41
Score 3	0.06	-0.29	1.00	0.09
Score 4	0.79	-0.41	0.09	1.00

The time series of scores of each mode are compared against the mean wind components of each wind region in Figure 4.13. This allows for illustration of the meaning of the time evolution of the principal components (scores) in relation to changes in the zonal and meridional components in each region. The changes in the scores time series for each mode match in some regions with those of the zonal wind component, or with

those of the meridional one or with both. A common feature is that the four subregions present a very similar evolution of their mean meridional component revealing that the variability of this component is rather uniform over all the CFN. However, the mean zonal component shows different variability and is the actual factor which seems to play a role in distinguishing subregions. The similarity of the time series of scores with the zonal or meridional component can be understood if the orientation of vectors in the loading maps is taken into account (Fig. 4.10). For instance, the EV subregion shows a similar evolution of the time series of scores and both the series of the zonal (r = -0.93) and meridional (r = 0.99) wind (Fig. 4.13a). This similar variability of both wind components seems to be associated with the NW-SE orientation of the Ebro valley and the channeling along it (Fig. 4.10a). This is also apparent in the forth subregion (NV, Fig. 4.13d) that also presents NW-SE orientation of the valleys (Fig. 4.10d) with the meridional and zonal components similar to the score (r = -0.84 and r = 0.77 respectively). The MS subregion evolves in agreement (r = 0.94) with the mean zonal component of the subregion (Fig. 4.13b) in concordance with the zonal orientation of the vectors in its loading map (Fig. 4.10b). Finally, the scores in NS subregion present also similar changes (r = -0.66) to the zonal mean wind component (Fig. 4.13c) due to the zonal orientation of the vectors with highest loads at the stations that define the subregion in its loading map (Fig. 4.10c).

Normalized spectra of the rotated scores provide a complementary understanding of wind temporal variability in each subregion. The series of scores are not continuous in time since the input data presented missing values and only daily fields with a high percentage of available measurements were selected (see section 4.1). Standard autocovariance Fourier transform spectrum analysis (Bloomfield, 1976) can find difficulties in its application in cases of large amounts of missing values. This case can be treated as one of irregular sampling of data (Belserene, 1988) and hence, spectra are calculated herein with an alternative approach that does not require equidistant sampling (Deeming, 1975). The spectral estimate is comparable to a normalized periodogram and can be interpreted as such, but no limitation is imposed on the regular or irregular character of the sampling when calculating the discrete Fourier transform. An spectral window is selected that contains the timescales of interest and the spectral estimate is obtained for a set of trial frequencies that in this approach, will be regularly distributed over the spectral window. For details and discussion on this approach the reader is referred to Deeming (1975) and Belserene (1988).

The normalized spectra of the rotated scores is displayed in Figure 4.14. The four subregions show wide spectral bands at low frequencies although they are not significant in comparison to a red noise autoregressive (AR1) process. It is worth noting that the time span of the scores series displayed in Fig. 4.13 is only of about three years (1999 to 2002) and thus of limited extension to significantly resolve the annual cycle. The low frequency bands accumulate the largest portion of variance in all subregions. In addition, the various subregions accumulate variance over specific intervals centered at periods which coincide with harmonics of the annual cycle. However few of these are of significance in comparison to an AR1. This is the case of the EV (region 1) and the NS (region 3) which display significant portions of variance at frequencies ranging between 2 and 4 months. Spectra of the EV and the NV subregions (Fig. 4.14a and d respectively)

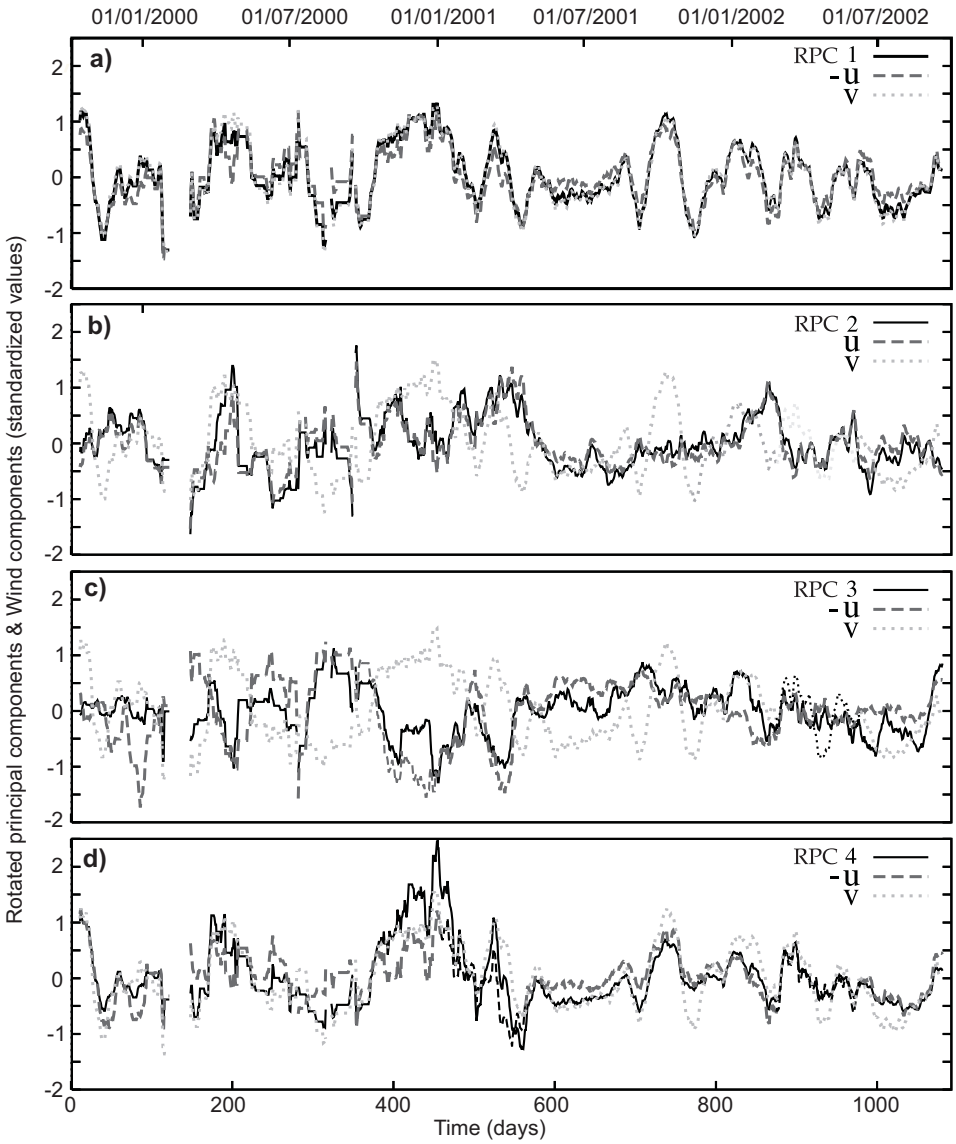

Fig. 4.13: 20 days moving average filter outputs of the time series of scores (solid lines) and the corresponding filtered u (dashed) and v (dotted) standardized mean wind components of the corresponding wind region as defined by Figure 4.11a: a) Ebro valley (region 1) b) mountain stations (region 2) c) North-to-South oriented stations (region 3) and d) northern valleys (region 4).

turn out to be very similar as should be expected from their correlation (Table 4.1). However, the northern valleys sites do not seem to receive AR1-significant contributions at high frequencies.

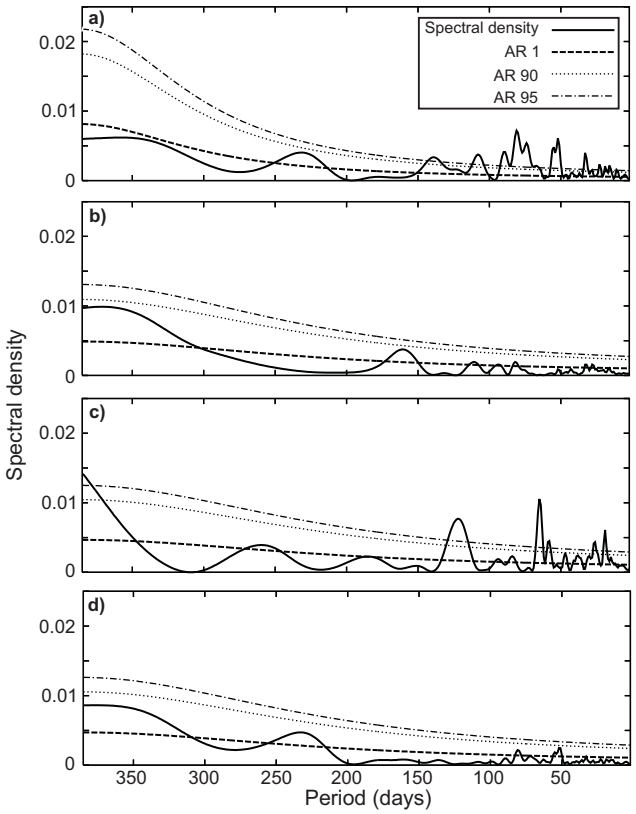

Fig. 4.14: Sample power spectra of the four principal component scores after VARIMAX rotation (solid lines), their first-order auto regressive process spectra (dashed lines) and the 90 (dotted lines) and 95% (dashed-dotted lines) confidence limits. The spectra represents the wind variability in each wind region: a) Ebro valley (region 1) b) Mountain stations (region 2) c) North-to-South oriented stations (region 3) and d) northern valleys (region 4).

For a more complete analysis of wind variability at low frequencies, the standardized anomalies of daily wind time series from the original extended dataset can be projected onto the eigenvector of each subregion in order to reach longer time series than those in the rotated scores (769 days). Projections were performed for the days with more than 50% of the measurements available, and thus a total of 2169 days were used. This set spans over a period of more than 7 years, since February 1995 up to September 2002. For each subregion, both the original scores and the time series obtained from projection are

very similar in the overlapping parts but the latter cover a longer time span allowing for increased spectral resolution at lower frequencies. Spectra for the projected time series are displayed in Figure 4.15. Overall, they show a similar behavior as the score spectra (Fig. 4.14), with a gain in resolution and suggesting the presence of significant variability at yearly time scales relative to an AR1 process. As in the case of Fig. 4.14 only the sites in the EV subregion (Fig. 4.14a) and the NS subregion (Fig. 4.14c) show significant contributions at higher frequencies. Figs. 4.14 and 4.15 suggest that the main difference in the variability of the first (EV) and forth (NV) subregion is the apparent lack of significant contributions to high frequency variability in the latter.

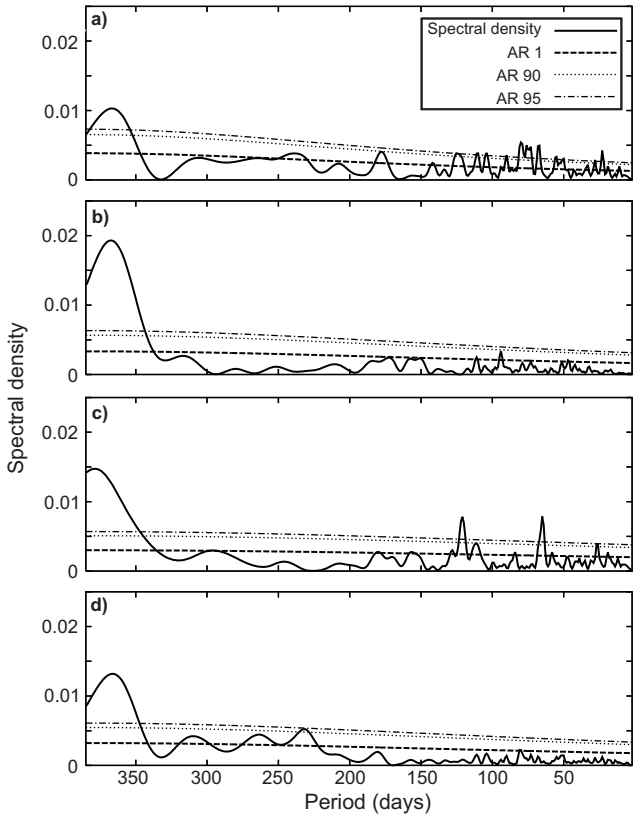

Fig. 4.15: Same as than Figure 4.14 but for the standardized projections of the daily wind fields (with more than 50 % of the data available) over the eigenvectors.

The climatological monthly means of the standardized proyections allow for the analysis of the annual evolution of the wind at the different subregions. These climatological means are displayed in its physical phase, in phase with the the zonal wind components of their subregions (Fig. 4.13), in Figure 4.16. As it could be expected, the EV subre-

gion shows a similar evolution as the NV subregion. The proyections, and therefore the zonal wind component, show the maximum in Summer (and February) and a minimum in Winter. The MS shows a similar behavior as the EV and NV during Winter but with opposite phase in the proyection (zonal wind component) during the Summer. Finally, the NS subregion shows the maxima in February and the minima in September.

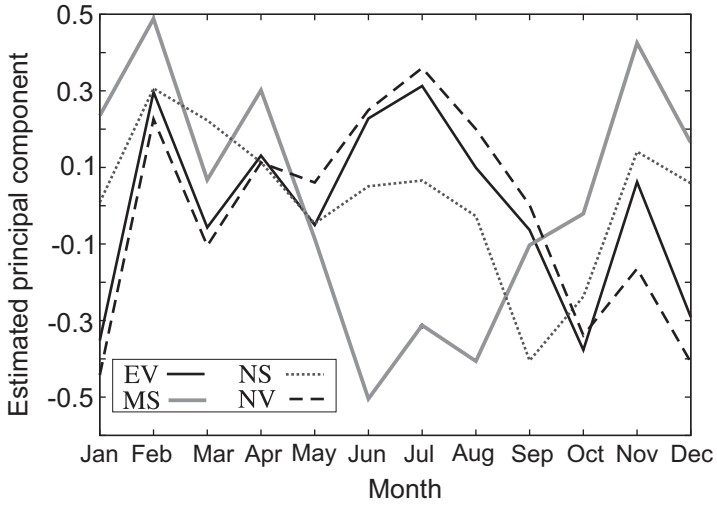

Fig. 4.16: Climatological monthly means of the standardized projections of the daily wind fields (with more than 50 % of the data available) over the eigenvectors.

4.5 Conclusions

In this chapter, two regionalization methodologies based on PCA were employed to classify together those observational sites with similar time variability. The first approach groups together stations with similar loads by CA producing a hard regionalization where the stations belong to only one subregion (Fig. 4.8c). The second approach rotates the principal modes and forms subregions allowing for some degree of overlapping (Fig. 4.11a). Each methodology presents its own advantages and drawbacks and the use of both schemes allows for a good degree of robustness in conclusions: both produce an equivalent array of groups providing wind regions in accordance with the topographic features of the terrain. The main difference between the two methodologies is that the one based on CA generates an area with few small subregions that are joined by the rotating methodology. The CA tends to form very specific subregions while the one based on the rotation forms more generic groups. However, this regionalization that rotates the principal modes, leaves two stations unclassified. Results from both classification will be

used in the next Chapter 5 to evaluate the ability of the WRF dynamical downscaling in reproducing the surface wind variability.

The variability of the meridional wind is very similar in all the subgroups and it is the zonal wind component the one that mostly contributes to characterizing the wind variability in each subregion. On the basis of this statement, it can be argued that an scalar PCA based approach applied to the zonal component would lead to similar results. While this is essentially true, such an approach would exclude the vectorial information and thus the possibility of highlighting patterns of distinct wind direction over the CFN (e.g. channeling along the valleys). The wind spectrum of each subregion was analyzed revealing the dominance of the annual cycle in all of them. Two subregions display significant variability at higher frequencies (2 to 4 months) in comparison with an AR1 process.

5

Regionalization in the WRF simulation*

The wind variability over the CFN has been analyzed with observations in the previous Chapter 4. The present chapter is devoted to analyze the capability of the WRF dynamical downscaling (Chapter 3) to reproduce the observed behavior. The evaluation focuses on the ability of the downscaling to reproduce the wind regions identified with observations in the previous chapter. It will be shown that the comparison of observations and the simulation at the regional scale provides a better agreement than the more frequent comparison at the sites. The realism displayed by the simulation during the evaluation allows for the use of the higher spatio-temporal coverage of the wind estimations to infer the wind variability where and when observations are not available.

5.1 Problematic and and rationale

As it was already discussed in Chapter 1, the evaluation of numerical simulations meets some uncertainties when comparison with local information is involved (von Storch, 1995). A frequent approach consist in the comparison of the observations against the nearest simulated grid point (Cox et al., 1998; Hanna and Yang, 2001; Buckley, 2004). However, this comparison is problematic since observations contains the influence of local effects that are not modeled and the simulation introduces *representativeness errors*. A more appropriate framework consist in compare observations and simulations at the regional scale (Reid and Turner, 2001).

This chapter evaluates in a first step the accuracy of the mesoscale model simulation to reproduce the observed wind variability by comparing the simulated and observed time series at the subregions identified with observations (Chapter 4). It will be illustrated how the use of average regional series damps local effects in the observations (noise)

* The main contents of this chapter are included in:
Jiménez, P. A., J. F. González-Rouco, E. García-Bustamante, J. Navarro, J. P. Montávez, J. Vilà-Guerau de Arellano, J. Dudhia, A. Roldan, 2009: Surface wind regionalization over complex terrain: evaluation and analysis of a high resolution WRF numerical simulation *J. Appl. Meteor. & Climatol.* (In Press).

as well as random *representativeness errors* in the simulation and thus, provides a first and suitable framework to evaluate the numerical simulation performance. Additionally, a complementary evaluation is obtained by applying the regionalization methodologies to the simulated wind and comparing results with those obtained in the observational analysis. This step illustrates the realism of the WRF simulation in reproducing the subregions found in the observational assessment. In order to provide comparable classifications, the regionalization methodologies are applied within the WRF simulation to a dataset that replicates in detail the temporal and spatial availability of observations along the restricted 1999 to 2002 period used to identify the subregions. For this purpose, the simulation is masked by introducing missing data, thus reproducing the spatial and temporal sampling of the observations. This allows to evaluate the model performance over a dataset that is most comparable to the one used in the observational study.

As discussed in Chapter 3, the distribution of wind sensors over a region of interest and the availability of observations through time is an issue of relevance that conditions our understanding of the variability of any meteorological observable. In order to investigate the impact that the lack of data over some subregions of the CFN can have on the analysis of surface wind variability, the WRF simulation is used as a tool to estimate the likely behavior of the wind field in areas and time intervals with scarcity of observations. This inference analysis will evidence the advantages of the higher spatio-temporal coverage that a numerical simulation provides in comparison with observational records. The analysis is accomplished in two steps, first by extending the wind regionalization to the complete simulated period (i.e. 1992 - 2005 instead of 1999 - 2002); and secondly, by considering the full simulated domain over the CFN (8640 grid points) over the whole period instead of only the grid points co-located with available observational stations. This allows to evaluate to what extent the limited temporal and spatial coverage of the observational network is appropriate to represent the wind variability over the area. Therefore, the results of this analysis are not only informative in assessing the realism of a long WRF simulation at high horizontal resolution in reproducing regional wind variability over an area of complex terrain, but also bear implications concerning the design of observational networks.

The rationale of the data preparation and the strategy used in the evaluation and inference analyses are described in Sections 5.1.1 and 5.1.2, respectively.

5.1.1 Evaluation of the WRF numerical simulation

The performance of the WRF simulation in reproducing the surface wind variability is evaluated from the perspective of two independent analysis. The first one evaluates the accuracy of the simulation to reproduce the temporal wind variability at the subregions previously identified in the observational data (Chapter 4). The second approach applies the regionalization methodologies to the simulated wind. This second evaluation not only allows to analyze the accuracy of the simulation in reproducing the wind behavior at the subregions (temporal wind variability), but it also allows for the evaluation of the realism of the simulation to identify the groups of stations obtained with observations (spatial wind variability).

Since the WRF simulation provides a higher spatial and temporal resolution than the observational network, it becomes necessary to manipulate the simulation to provide comparable datasets. In particular, two datasets are created by transforming the simulation in order to replicate the spatial and temporal distribution of wind observations at the 35 observational sites (Fig. 4.1):

- The spatially and temporally masked simulation (STMS): the simulated wind at each one of the 35 observational sites is represented by the meridional and zonal wind components at the nearest co-located grid points. Here, spatial masking refers to the fact that only the grid points co-located to the observational sites are used whereas the temporal mask means that the simulated values corresponding with a missing observation are suppressed. This dataset spans the complete observational period from 1992 to 2005.
- The reduced spatially and temporally masked simulation (RSTMS): this dataset consists of the 769 days used to identify the observational subregions (Section 4.1). The subset of days spans the 1999 to 2002 period and are extracted from the STMS dataset. This dataset will be referred to as the RSTMS in allusion to its equivalence with the reduced observational dataset.

The ability of the numerical simulation in reproducing the wind variability over the CFN will be firstly evaluated by comparing the observed and simulated mean wind at the different subregions identified with observations. The regional time series are calculated within the complete observational dataset and its counterpart from the simulation, the STMS subset, in order to use comparable datasets covering the whole observational period. This approach allows for an evaluation of the accuracy of the simulation in the regions obtained in the observational analysis. However, the question remains open as to whether the model would provide a similar classification of subregions if the same regionalization analysis performed on observations would be carried out within the WRF simulation.

Therefore, a complementary evaluation of the WRF simulation accuracy to reproduce the observed wind variability will be attained by applying the two different regionalization approaches to the RSTMS simulated surface wind and comparing results with those achieved with observations (Figs. 4.8c and 4.11a). It should be noted that, in spite of relatively high resolution, such comparison is still subjected to the effects of local features that contribute to variance in the observations and that are not captured by the model; in addition, on the side of the WRF simulation the wind is affected by *representativeness errors*. Therefore, a detailed agreement between the resulting regionalizations is not expected in this comparison, nevertheless, similarities on their main characteristics of simulated and observed climate should be expected.

5.1.2 Inference analyses: on the effects of observational sampling

On the basis of the accuracy displayed by the simulation during the evaluation phase, the simulation is subsequently used to address the impact that the limited temporal and spatial coverage of observations may produce on the analysis of the wind variability

over the CFN. With this aim, the wind regionalization methodologies are applied to the simulated wind from the innermost domain (Fig. 3.1) avoiding temporal and/or spatial masking. Additional versions of the dataset are created to adapt the simulated wind to the specific purposes of the inference analysis:

- The spatially masked simulation (SMS): this dataset consists of the simulated wind at each one of the 35 observational sites represented by the meridional and zonal wind components at the nearest co-located grid points. It does not include any temporal masking and therefore the SMS dataset spans the whole simulated period from 1992 to 2005 without any missing value.
- The complete simulation: this dataset consists of the wind time series at all the simulated grid points from the innermost domain (8640 points) over the whole simulated period from 1992 to 2005. Therefore, this dataset contains all available information from the WRF numerical simulation within the area of study.

The regionalization obtained with the SMS dataset will be compared with the regionalization obtained with observations for the 1999 to 2002 period (Figs. 4.8c and 4.11a). This comparison allows for analyzing the influence of the length of the observational period in identifying the subregions. The inference study will be extended in a second step, in which the regionalization is obtained involving the complete simulation. The comparison of the resulting wind regions against the ones obtained previously with the SMS dataset allows for assessing the influence of the observational spatial coverage in identifying the subregions. These two analyses allow to infer on the characteristics of wind variability outside the temporal and spatial domains covered by observational sampling. From this point of view the study also bears implications for the design of observational networks.

The temporal coverage and the number of time series in each observational (reduced and complete) and simulated dataset (STMS, RSTMS, SMS and complete) used in the study are summarized in Table 5.1 to aid the understanding of the experimental set up and the explanation of the results.

Table 5.1: Main characteristics of the wind datasets used in the comparison of observations and the WRF simulation: temporal coverage, number of time series and a flag indicating the presence or absence of missing values.

Temporal coverage	time series	Missing values	Observational dataset	Simulated dataset
1999 - 2002	35	yes	Reduced	RSTMS
1992 - 2005	35	yes	Complete	STMS
1992 - 2005	35	no	——	SMS
1992 - 2005	8640	no	——	Complete

5.2 Local versus regional comparisons

Before evaluating the performance of the WRF simulation at the regional scale, it is interesting to analyze its behavior at the local scale. This also helps to illustrate the advantages of the regional approach adopted here and the added value of the regional simulation respect to the original reanalysis fields.

The analysis is based on the 769 days of the reduced dataset used to identify the subregions, and its equivalent dataset from the WRF simulation, the RSTMS (Table 5.1). The comparison of both datasets is displayed in Figure 5.1, where the correlation and RMSE between observed and simulated wind are plotted for each site and wind component. Their values are depicted with circles (correlation) and circumferences (RMSE). The RMSE at each site is calculated with the time series of the anomalies, after removing the climatological mean, and subsequent normalization by the standard deviation of the observational time series. The radius of the circles in the correlation panels is proportional to the correlation score between the observational time series and the simulation at the nearest grid point, whereas the radius of the external circumferences indicates the highest correlation score among the observations at a given site and the nearest 25 grid points. For the RMSE panels, the radius of the external circumferences is proportional to the RMSE between the observational time series and the nearest simulated grid point, whereas the radius of the internal circles are proportional to the smallest RMSE among the observations and the nearest 25 grid points. For the case of the RMSE, the circle is colored in white (gray) to indicate that the RMSE is lower (higher) than one. In addition, the external circumference is filled up in gray at those sites wherein the lowest RMSE among the nearest 25 grid points is smaller than one and the RMSE calculated with the nearest grid point higher than one.

The zonal component shows an increasing gradient of the correlation to the South of the CFN, with the highest values at stations located in the EV and the MS subregions (Fig. 5.1a). At the lowest correlated sites, mainly located at the NS and NV subregions, the normalized RMSE shows the highest values (often above one, Fig. 5.1b), indicating a degradation in the accuracy to reproduce the temporal variability at these locations. The meridional wind component shows high correlation values (Fig. 5.1c) except at stations 15 and 26 (Fig. 4.1). As a consequence, these sites present high normalized RMSE values (Fig. 5.1d), indicating a low accuracy of the model to reproduce their temporal variability. A few locations mostly located in the NV subregion show RMSE values higher than 1.0 but generally, the RMSE presents values below 1.0 (Fig. 5.1d).

As it was explained in the previous section, the comparison of observations and simulations at the local scale is hampered by uncertainties associated with local effects and *representativeness errors*. Local effects are intrinsic to observations and a quantification of their influence on the comparison is not straightforward. However, a quantification of the *representativeness error* can be obtained by comparing the site-observations with the nearest and more distant grid points. With this aim, the highest correlation between the site and the 25 nearest grid points is also displayed in Figure 5.1. This value is represented with a circumference, and it is only visible when a correlation value larger than that corresponding to the closest grid point, represented with a circle, is found. In the case of the RMSE, the smallest value among site observations and the nearest 25

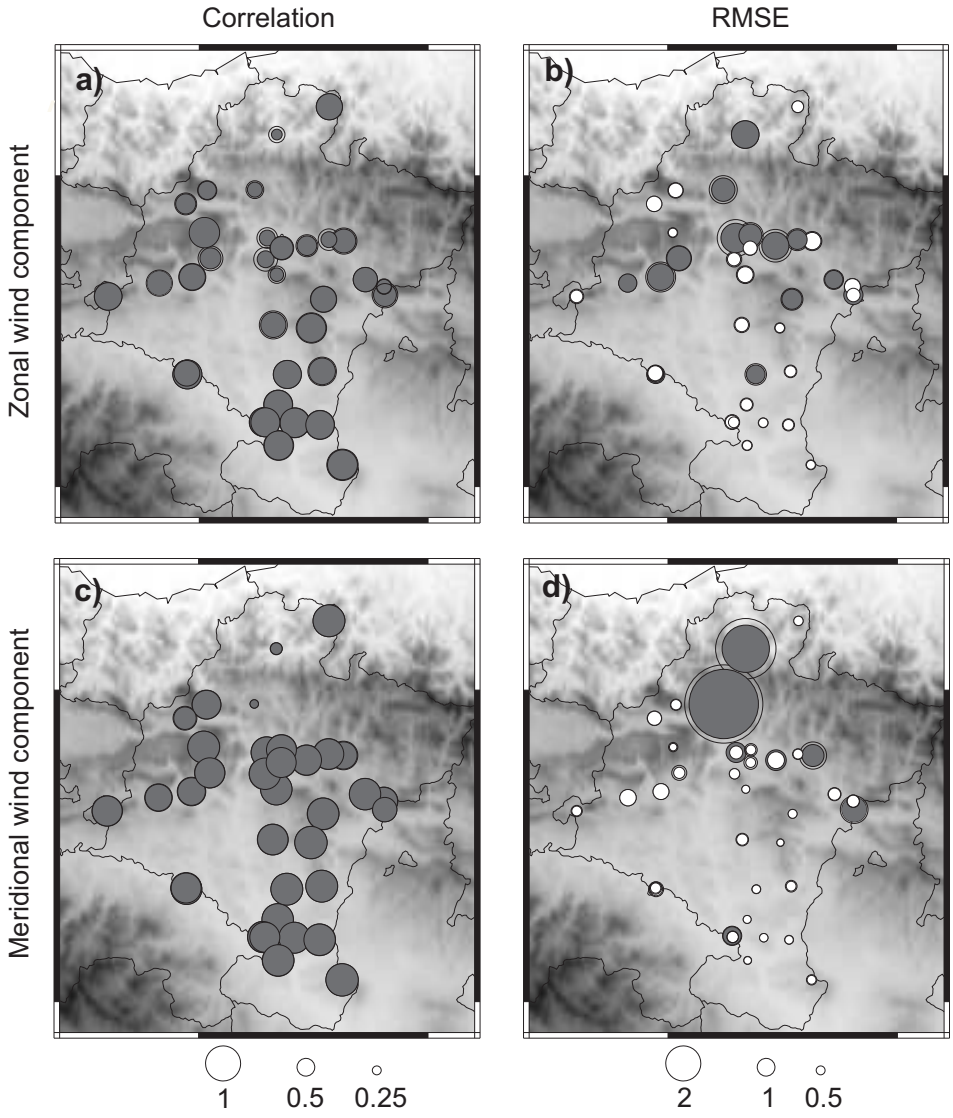

Fig. 5.1: Correlation (left panel) and RMSE (right panel) of the observed and simulated zonal (first row) and meridional (second row) wind components calculated with the reduced datasets that spans the 1999 to 2002 period (Table 5.1). See the text for details.

grid points is represented with a circle. The circle can be distinguished from its external circumference that represents the RMSE obtained with the co-located grid point only when a smaller RMSE is found with more distant grid points.

The sites with the best RMSE and correlation values do not show any improvement when a grid-point different from the nearest co-located one is considered (Fig. 5.1). At the sites with the poorer scores, mostly located at the NS and NV subregions, there is a tendency to show better correlation and RMSE at some nearby grid point than at the closest one. This is an indication that a certain misrepresentation of the observed variability could be taking place by assigning observations to the nearest grid point. However, the effects of the *representativeness errors* do not seem to produce a drastic change on the simulation performance. Hence, the representation of the site-observations by its co-located grid point seems to be a reasonable approximation for this particular case.

Another argument that has also been used in the previous section is that the average of local observations into regional series filters out local variability (noise) and enhances the regional signal. Since Section 5.3.1 will describe an evaluation of the model performance at the scale of the subregions found with observations, it is interesting to illustrate the impact on the simulation performance associated with the change from local scales to subregional averages. Figure 5.2 shows Taylor diagrams (Taylor, 2001) calculated with the simulated and observed wind components at each one of the sites belonging to a given wind region and their regional series. Again, the analysis is restricted to the 769 days of the reduced datasets (Table 5.1). In most of the subregions, the average series are in a better agreement than when comparisons are made at the site level. A clear example is the zonal wind component of the MS subregion (Fig. 5.2c) or the meridional wind component of the NS subregion (Fig. 5.2f). One would have expected to find the point associated with the regional comparison around the centroid of the cloud of points comparing local information. However, the point associated with the regional evaluation is displaced towards the (1,0) point that represents a perfect agreement in the Taylor diagrams. This result shows the beneficial effect of averaging the observed and simulated series to mitigate local effects and *representativeness errors* in order to reinforce the regional signal, and it is the basis of the evaluation assessment performed in this work.

The decision of evaluating at the regional scale instead of at the local scale leads to question whether it is necessary or not to simulate at such a high spatial resolution as the one used herein (2 km). On one hand, increasing the horizontal resolution provides a better representation of the orography, a factor that potentially contributes to a better simulation of local variability (Fig. 5.2). On the other hand, there are some locations at which the performance of the simulation is rather low in spite of the high horizontal resolution used (e.g. stations 15 and 26). The lack of realism at some places poses the question of whether it would be possible to reduce the horizontal resolution without degrading in excess the overall simulation behavior and still keep a satisfactory regional evaluation. Addressing this topic probably requires performing sensitivity studies with additional simulations at different horizontal resolutions. Such an analysis is beyond the purposes of the present study.

As a final illustration of the performance of the WRF simulation at different spatial scales, it is interesting to consider the added value of the WRF simulation with respect to

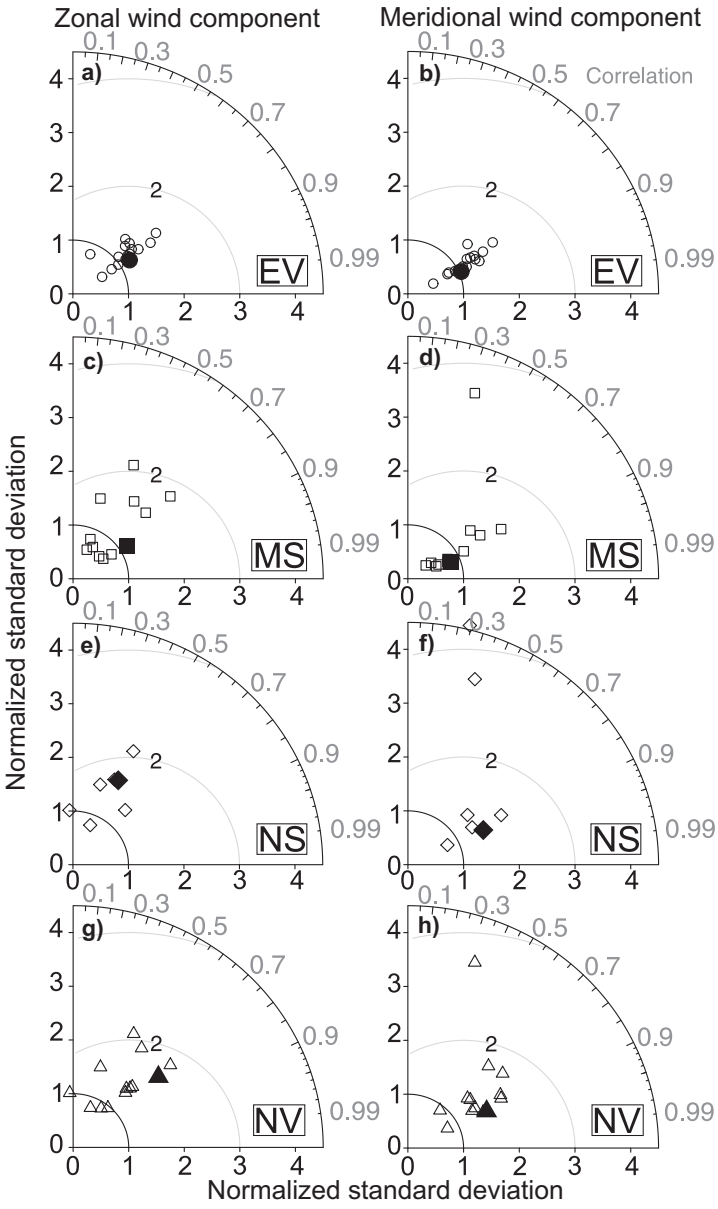

Fig. 5.2: Normalized Taylor diagrams comparing the zonal (left panels) and meridional (right panels) wind components at each one of the four subregions identified with observations (Fig. 4.11a). The hollow symbols represent the comparison at each observational site whereas the solid symbols the comparison of the regional time series. The reduced datasets are used in the comparison (Table 5.1).

the ECMWF wind data used as initial and boundary conditions. This point is analyzed by comparing the WRF simulated wind and the ECMWF surface data against observations. For this purpose, Taylor diagrams calculated with the observed and simulated (WRF and ECMWF) wind components at the four subregions are analyzed. In the case of the WRF simulation, the regional wind components are calculated with the RSTMS (Table 5.1) as in Figure 5.2. In the case of the ECMWF, the surface wind components at the four grid points closest to the CFN (Fig. 4.1) are compared with the regional wind components at each subregion. It should be noticed that four ECMWF wind series are compared against the mean observed wind at each subregion, whereas a single simulated series for each subregion is used in the case of the WRF model. This is a consequence of the impossibility of obtaining a time series for each subregion with the relatively coarse spatial resolution of the ECMWF data (1^o) and is at the core of the improvement introduced by the dynamical downscaling.

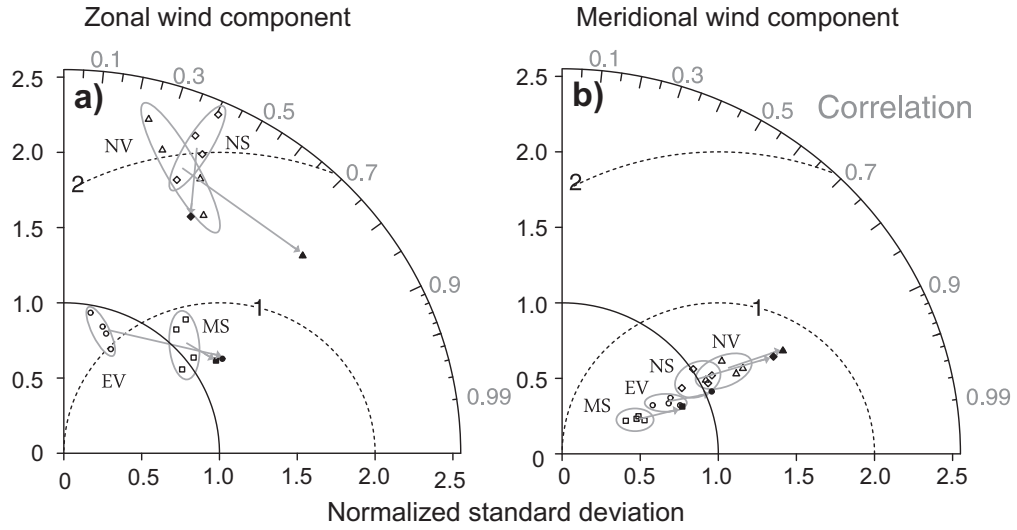

Fig. 5.3: Normalized Taylor diagrams comparing the daily mean zonal (a) and meridional (b) wind components from the WRF simulation (solid symbols) and the ECMWF reanalysis/analysis (hollow symbols) against the observed wind at each one of the four subregions (Fig. 4.11a): EV (circles), MS (squares), NS (diamonds) and NV (triangles). In the case of the WRF simulation, the comparison is performed for the 769 days of the reduced datasets (Table 5.1). For the ECMWF, wind data at the four grid points closest to the CFN (Fig. 4.1) are used in the comparison. The arrows highlight the different performance displayed by the ECMWF and the WRF simulation.

Results are displayed in Figure 5.3. A better skill in reproducing the meridional wind component variability (Fig. 5.3b) than the zonal one (Fig. 5.3a) can be appreciated

for both ECMWF and WRF simulations. This is not surprising in view of the results displayed in Figures 5.1 and 5.2. Notice that the solid symbols denoting the WRF performance at the regional scale in Figure 5.2 and 5.3 represent the same information. The better reproduction of the meridional component may be interpreted as a consequence of the specific location of the CFN between the two large mountain systems in the north of the Iberian Peninsula (i.e. the Cantabrian system and the Pyrenees, Fig. 1.2) that strongly modify the surface circulations blocking and channeling the flow between them and probably being responsible of the similar meridional wind variability observed over the whole CFN. The mountain systems are large enough to be represented in the simulation even with the coarser ECMWF spatial resolution (more than 100 km), and therefore the reanalysis data capture the meridional wind variability introduced over the CFN as a consequence of the surface flow channeling. For the meridional component (Fig. 5.3b), correlation values between WRF or ECMWF and observations are around 0.9 and normalized values of RMSE below 1.0. The WRF simulation slightly increases the correlation shown by the ECMWF data (Fig. 5.3b) displaying a tendency to increase the variability which leads to an overestimation in the NS and NV subregions; yet, it is in better agreement with the observations in the MS and EV subregions. The WRF simulation shows a clear improvement for the zonal component (Fig. 5.3a), especially for the EV and NV subregions wherein the correlation increased from 0.2 to 0.85 and from 0.4 to 0.8, respectively. The best agreement between ECMWF data and observations in the case of the zonal component is found for the MS subregion. Since the zonal wind variability causes the differences between the subregions, it can be expected that the inclusion of a more realistic orography will substantially impact the accuracy of its simulation. In this way, the coarser ECMWF horizontal resolution is not able to resolve the CFN mountain ridges leading to low correlation at the valley subregions and better scores for the MS subregion. In turn, the higher spatial resolution of the WRF simulation improves the representation of the terrain features leading to an improvement in reproducing the surface zonal wind component variability at the valley subregions.

5.3 Evaluation: WRF replication of observed regional variability

The evaluation of the WRF performance to reproduce the regional wind variability is accomplished from two slightly different perspectives: 1) comparison of the temporal wind variability at the subregions identified with observations in Chapter 4, and 2) application of the regionalization methodologies to the simulated wind and comparison with those obtained with observations.

5.3.1 Evaluation over observational regions

The correlation, bias and normalized RMSE between the regional wind components calculated with the complete observational dataset and the STMS dataset (Table 5.1) are displayed in Table 5.2. Hence, the complete period of availability of information (1992 to 2005) is used in the comparison. A good concordance is found between these scores and the ones displayed in Figure 5.3 corresponding to the reduced datasets (1999 to 2002).

This extends the previous interpretations to the whole observational period. To further illustrate this statement, the regional time series of the zonal component from the STMS and the complete observational datasets (Table 5.1) are displayed in Figure 5.4. This component has been selected for illustration since it is responsible for the different wind variability of the subregions. The temporal period covered by the reduced datasets (Table 5.1) is highlighted in gray. The series show similar temporal variability within and outside this period in concordance with the similar scores found for the reduced and complete periods (Fig. 5.3 and Table 5.2, respectively). The time evolution of the RMSE is also displayed in Figure 5.4. There is a tendency to show lower RMSE values during summer than winter revealing a certain seasonal behavior in the performance of the numerical simulation. The varying number of available stations within each region (Fig. 5.4) does not seem to impact the quality of the simulation in the EV, MS and NV subregions, but it could be of relevance for the NS subregion where the number of available sites is low, particularly after year 2000.

Table 5.2: Correlation, Bias (observations minus simulations) and RMSE of the anomalies (normalized by the standard deviation of the observations) calculated with the regional time series of the zonal/meridional wind components from the complete observational dataset and the STMS (Table 5.1).

	$Corr$ u/v	$Bias$ u/v	$RMSEa/Sd_{obs}$ u/v
EV	0.86 / 0.91	0.46 / -0.72	0.62 / 0.43
MS	0.86 / 0.91	-0.24 / -0.86	0.52 / 0.43
NS	0.54 / 0.86	-0.48 / -0.87	1.31 / 0.63
NV	0.74 / 0.86	-0.21 / -0.31	1.27 / 0.71

In order to avoid introducing potential biases by constructing regional time series from a variable number of stations (Fig. 5.4) with different variances, each time series was standardized before calculating the regional average. The correlation between these observed and simulated regional time series (r = 0.85, 0.85, 0.52 and 0.76 for the EV, MS, NS and NV, respectively) are very similar to those obtained previously with the averaged raw series (Fig. 5.4). This indicates that the regional signal is neither degraded nor improved with the standardization and therefore is not subject to the dominant variability of a few sites.

Complementary information to evaluate the performance of the simulation in reproducing the observed temporal variability can be obtained through a wavelet analysis (Foufoula-Georgiou and Kumar, 1995; Torrence and Compo, 1998). The wavelet power spectra of the simulated and observed time series must compare well to indicate a good performance. These are shown for the observed and simulated regional zonal component of the EV and MS subregions in Figure 5.5. The wavelet spectra of the MS subregion show two cones of influence as a consequence of the missing data (Fig. 5.5b,d). The EV shows high similarities between the observed and simulated power spectra (Fig. 5.5a and c, respectively). It shows activity at the highest frequencies during the whole observational

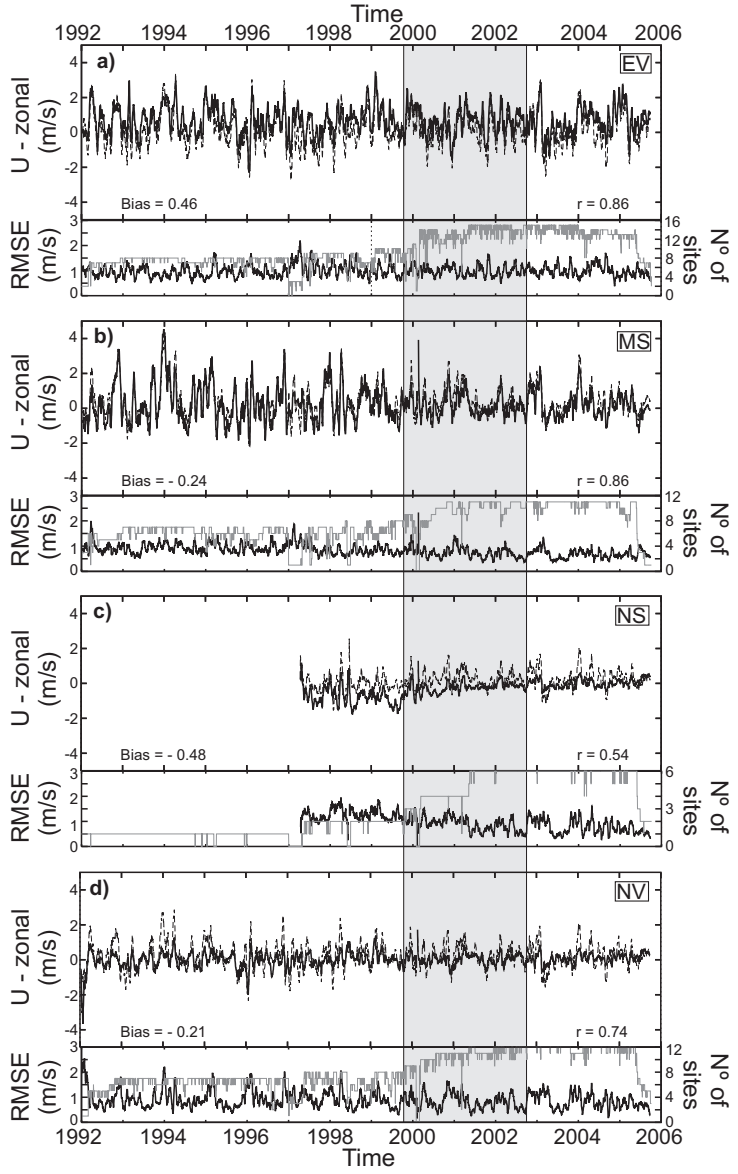

Fig. 5.4: 20 days moving average filter outputs of the zonal wind component at the a) EV, b) MS, c) NS and d) NV subregions (Fig. 4.11a) calculated with the complete observational dataset (solid) and the STMS (dashed). The correlation and the bias (observed minus simulated mean) of the regional series before applying the moving average are also displayed. The lower panels represent the 20 days moving average RMSE (black) and the number of available stations at each particular day (gray). The gray area highlighted on the time series represents the time span of the reduced datasets (Table 5.1).

period with an alternation of significant contributions at lower frequencies corresponding to the annual cycle. The MS, on the contrary, shows an observational wavelet spectra with a significant annual band during the whole observational period (Fig. 5.5b). The simulation reproduces an important spectral density around the annual band, although it is not significant during the whole period as in the observations (Fig. 5.5d). Also, the reduction in variance in the interval between 200 and 20 days that takes place at MS after year 2000 is captured by the simulation. This reduction could potentially be attributed to the variable number of stations available to calculate the regional time series (Fig. 5.4). In order to evaluate to what extent the wavelet power spectra is sensitive to this limitation, the analysis was performed using the complete time series from the numerical simulation (SMS, Table 5.1). The resulting wavelet power spectra are displayed for the EV and MS subregions in Figure 5.5e and f, respectively. Both spectra are similar to the ones calculated with a variable number of stations. This rules out potential effects on the regional variability associated with the variable number of stations. The wavelet power spectra of the zonal wind component at the NS and NV subregions are also in agreement with the simulation (not shown); nevertheless, a certain degradation becomes evident at the NS subregion as could be expected from a lower number of available time series (Fig. 5.4c) and a worse reproduction of the wind variability at two of its sites (Fig. 5.1).

Since the meridional wind variability is very homogeneous over the CFN, the regional wavelet power spectra are similar in all the subregions, and show a similar structure as the one displayed for the zonal wind of the EV subregion (Fig. 5.5a) wherein both wind components present equivalent variability (see Chapter 4). A higher agreement between the wavelet spectra calculated with observations and simulations is found for this component indicating the higher accuracy of the simulation in reproducing it (Table 5.2).

5.3.2 Regionalization in the WRF simulation

The previous evaluation is subjected to the constraint of employing the groups of stations obtained with observations (Fig. 4.11a) to determine the regional time series. In order to provide additional insight about the numerical simulation performance and avoid this limitation, the regionalization methodologies used to identify the wind subregions (see Section 4.2) are applied to the WRF simulation. The RSTMS dataset is used in order to involve a subset of data comparable to the reduced dataset used to identify the observed subregions (Table 5.1).

The first step in both regionalization methodologies is a PCA (Preisendorfer, 1988; von Storch and Zwiers, 1999). The regionalization based on the rotation of the main modes of variation identifies one subregion from each retained mode. Since the purpose of this analysis is to evaluate the accuracy of the numerical simulation to reproduce the spatial coverage and the wind variability of the subregions identified with observations (EV, MS, NS and NV), a total of four principal modes are retained. This decision is in agreement with results from the methodology based on CA for which the number of identified subregions does not depend on the number of principal modes retained, and four subregions appear as a reasonable number to retain.

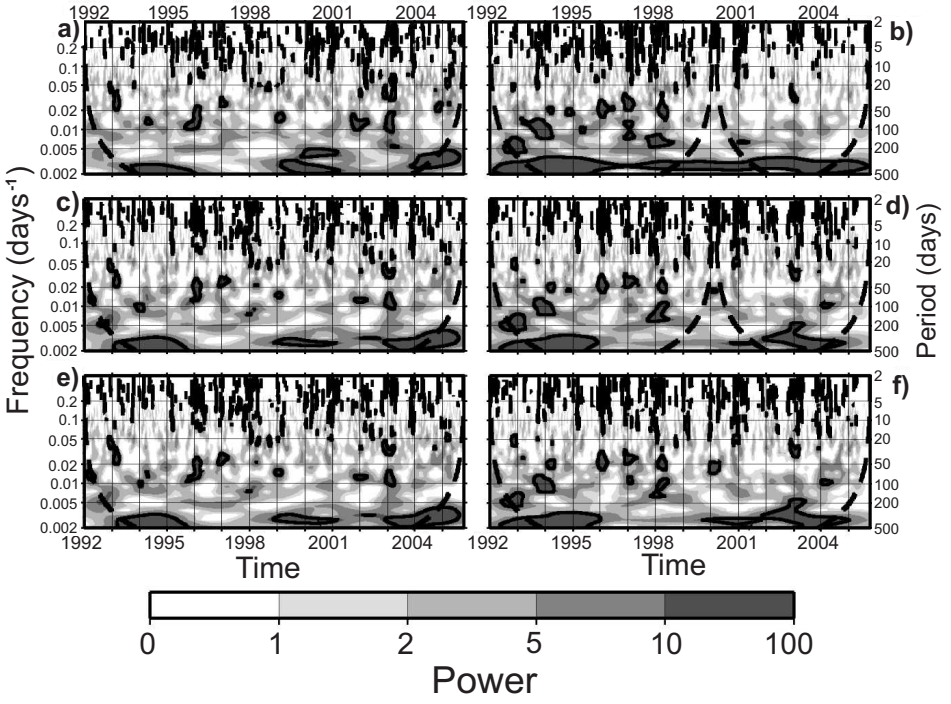

Fig. 5.5: Wavelet spectral power of the zonal wind component time series for the EV subregion (left panels) and the MS subregion (right panels) built from: the complete observational dataset (top), the STMS dataset (middle) and the SMS dataset (bottom). See Section 5.1 for details on the datasets. The shaded areas represent normalized variances higher than 1, 2, 5 and 10 whereas the black thick contour lines enclose regions of confidence above 95 % for a first-order autoregressive process. The dashed line represents the "cone of influence", below which edge effects become important (Torrence and Compo, 1998).

The classification obtained with the CA methodology is displayed in Figure 5.6a. A number of stations distributed over the Ebro valley (circles) are grouped together. This subregion is similar to the EV subregion identified with observations (Fig. 4.8c) but in this case the stations to the south of the *western mountains* are not included in the subregion. The MS subregion also seems to be discriminated and formed by stations in the *western* and *eastern mountains* (squares), but it does not include the mountain stations located in the *northern mountains* (Fig. 4.8c). A NV subregion is recognized (triangles), gathering together some stations as in its observational counterpart (Fig. 4.8c). Finally, a subregion formed by stations in the Northwest of the CFN (diamonds) is recognized close to the location the NS subregion (Fig. 4.8c). In summary, the regionalization obtained (Fig. 5.6a) is similar to that one obtained with observations (Fig. 4.8c) but some differences in the boundaries of the subregions are evidenced. These minor discrepancies were expected

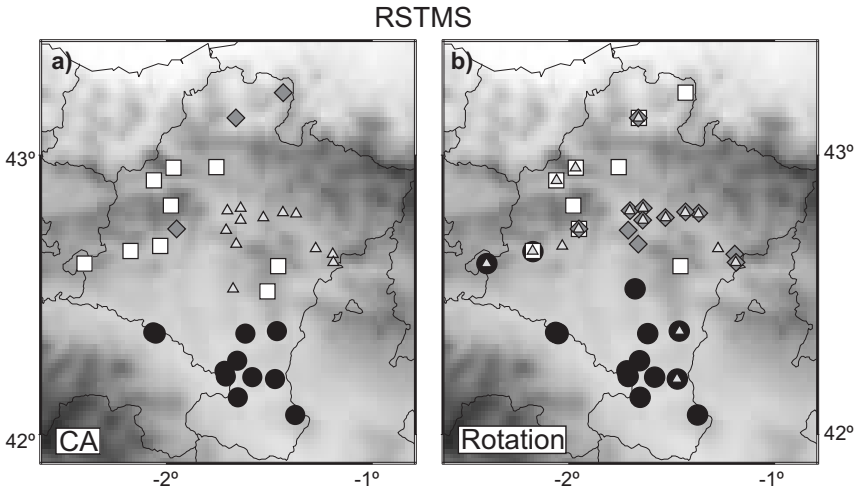

Fig. 5.6: Regionalization configurations attained after applying the methodology based on the CA (a) and the methodology based on the rotation of the principal components (b) to the RSTMS dataset (Table 5.1). The different symbols represent locations with similar temporal wind variability which define the identified subregions.

since the observations are affected by local effects that are not modeled and the simulation suffers from the *representativeness error*.

The resulting regionalization obtained with the rotational procedure can be observed in Figure 5.6b. The overlapping classification that this approach provides in contrast to the rigid CA-method becomes evident. The EV valley subregion (circles) is mainly formed by the same stations as with observations (Fig. 4.11a). The MS subregion (squares) is also identified without including a couple of mountain stations located in the *eastern mountains*. Another subregion (triangles) mainly associated with the NV is also identified, also including some stations located at the Ebro valley. A NS subregion seems to be clearly identified this time (diamonds) including a large number of stations along the NV subregion. Hence, a similar spatial structure as the regionalization with the observed surface wind is reproduced (Fig. 4.11a), the main difference being the enlargement of the NS subregions with stations located in the NV subregion. The problems in the identification of the NS subregion are not surprising if the poorer scores attained here are recalled (Fig. 5.3).

Since this last regionalization methodology assigns one principal mode of variation to each subregion, the rotated principal components provide information of the temporal wind variability at the subregions. Hence, the rotated principal components in the simulation can be compared against their counterparts in the observations to evaluate the performance of the numerical simulation in reproducing the temporal wind variability at the subregions. This comparison is depicted in Figure 5.7.

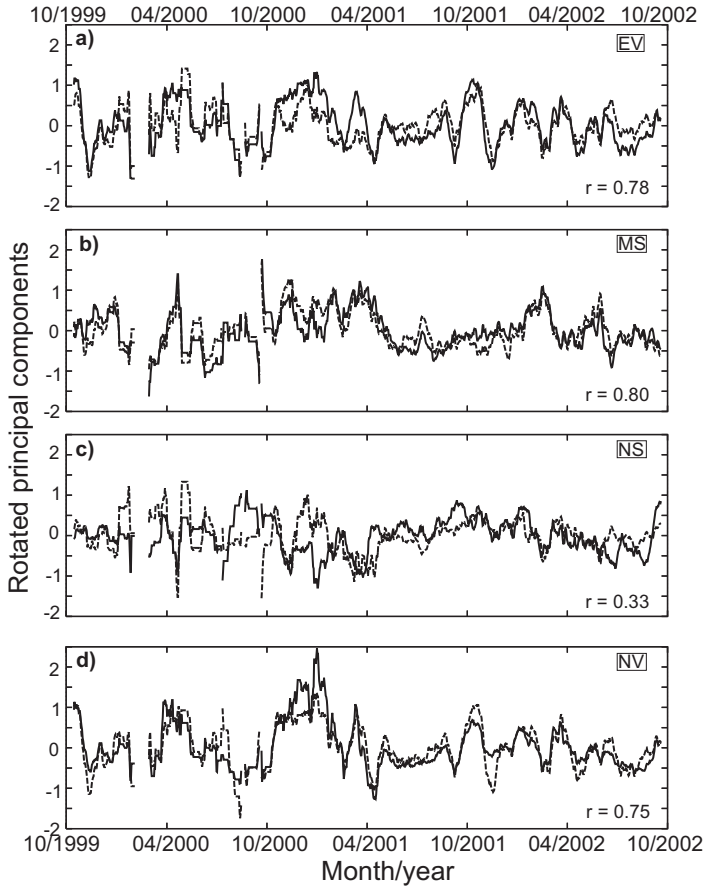

Fig. 5.7: 20 days moving average filter outputs of the rotated principal components associated with the a) EV, b) MS, c) NS and d) NV subregions. The solid lines represent the calculations performed with the reduced observational dataset (Table 5.1) whereas the dashed lines represent the RSTMS dataset (Table 5.1). The correlation between the unfiltered time series at each particular subregion is also shown.

The principal components associated with the EV subregion (Fig. 5.7a) present a high correlation (r = 0.78). In the analysis performed with observations, it was found that the EV principal component presented high correlation with both the mean meridional and zonal wind components of the subregion (r = 0.99 and -0.93, respectively) which is reproduced by the simulation although the scores are not so high (r = 0.73 and -0.86, respectively). The MS subregion also displays similar observed and simulated principal components (r = 0.80, Fig. 5.7b). The variability of this mode was found to be in agreement with the mean zonal wind component (r = 0.94); although the correlation

diminishes in the simulation, it still presents a high value (r = 0.80). The NS subregion shows a rather low correlation score among the observed and simulated principal components (r = 0.33). The agreement is better after year 2000 which could be associated with the higher number of stations available after this year (Fig. 5.4c). The rotated principal components associated with the NV subregion (Fig. 5.7d) show a high correlation (r = 0.75). The NV rotated principal component presents a high correlation value with the simulated mean zonal wind component of the subregion (r = -0.67) though not as high as the correlation found with observations (r = -0.84). On the contrary, this subregion displays a higher value for the meridional wind component (r = 0.89) if compared to that with observations (r = 0.77).

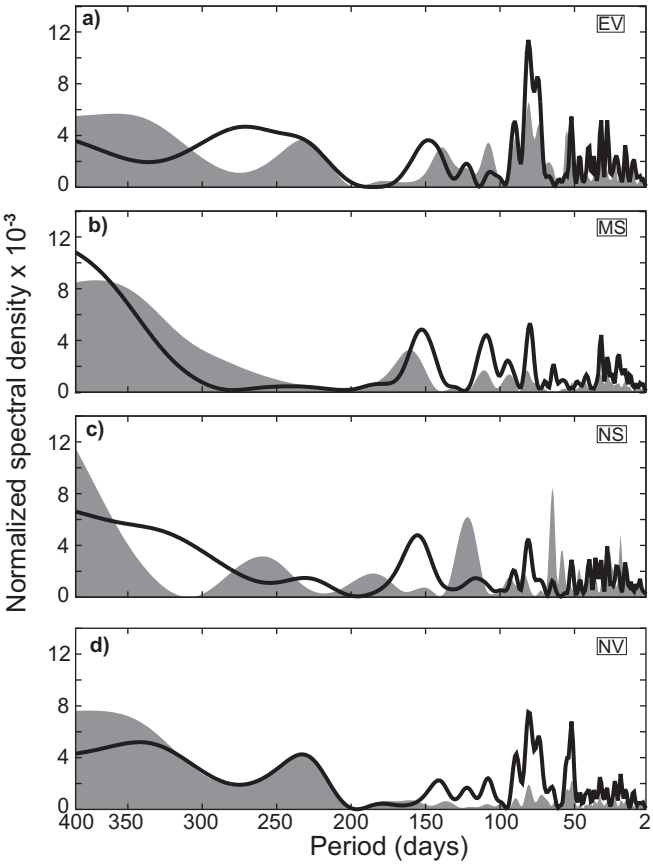

Fig. 5.8: Spectral analysis of the rotated principal components displayed in Figure 5.7. The shaded areas (black lines) represent the normalized spectra calculated with the observational (simulated) series.

Complementary information for the evaluation of the temporal wind variability at the subregions can be provided through spectral analysis of the rotated principal components. Since the RSTMS and observational datasets span only the 1999 to 2002 period and present a considerable number of missing data, the wavelet analysis is avoided in this case. Thus, information concerning the spectral wind variability is obtained in this case through analysis of the rotated principal components with techniques that allow for the presence of unevenly spaced data (Deeming, 1975; Belserene, 1988). See Chapter 4 for details on this methodology. Normalized spectra of the simulated and observed rotated principal components are displayed in Figure 5.8. The EV subregion shows a reasonably good replication of the observed peaks and bands, somewhat distorted at the lower frequencies (Fig. 5.8a). The MS spectrum is also quite similar to the observational one, although seems to overestimate the higher frequency variance, and the lowest frequency band is somewhat displaced with respect to observations (Fig. 5.8b). The NS subregion presents the most distorted spectrum (Fig. 5.8c) which constitutes a further evidence of its poorer wind simulation within this region (Figs. 5.3 and 5.7c). This could be attributed, at least in part, to the small number of stations that form this subregion and a worse reproduction of the wind variability at two of its sites (15 and 26, Figs. 4.1 and 5.1). The NV subregion shows qualitative agreement between the observational and simulated spectra with an overestimation of the spectral density at the highest frequencies while being closer to observations at the lower frequencies (Fig. 5.8d). It is interesting to notice the similar spectra displayed by the EV (Fig. 5.8a) and NV (Fig. 5.8d) as a consequence of the similarities of the wind variability at these subregions.

5.4 Inference analysis: the effects of the observational sampling

This subsection exploits the higher spatial and temporal sampling of the WRF simulation to further analyze the surface wind variability over the CFN. This analysis leans on the accuracy displayed by the simulation in identifying the wind subregions and reproducing the observed wind variability in the previous sections.

5.4.1 Regionalization over the 1992-2005 period

The simulated wind time series wherein no missing values are present can be used to analyze the influence of the observational period length used to identify the subregions with observations (1999-2002). This allows to answer the question of how the regionalization would be affected by extending the observational period. For this purpose the regionalization methodologies are applied to the SMS dataset, which consists of the complete time series from 1992 to 2005 of the simulated grid points co-located near the 35 observational sites (Table 5.1).

The classification obtained with the CA approach can be observed in Figure 5.9a. The resulting regionalization compares well with the output of this methodology for the shorter period (RSTMS, Fig. 5.6a). Hence, similar comments as the ones stated for the RSTMS classification can be made concerning the accuracy to reproduce the observational subregions (Fig. 4.8c). The approach based on rotation also displays similar

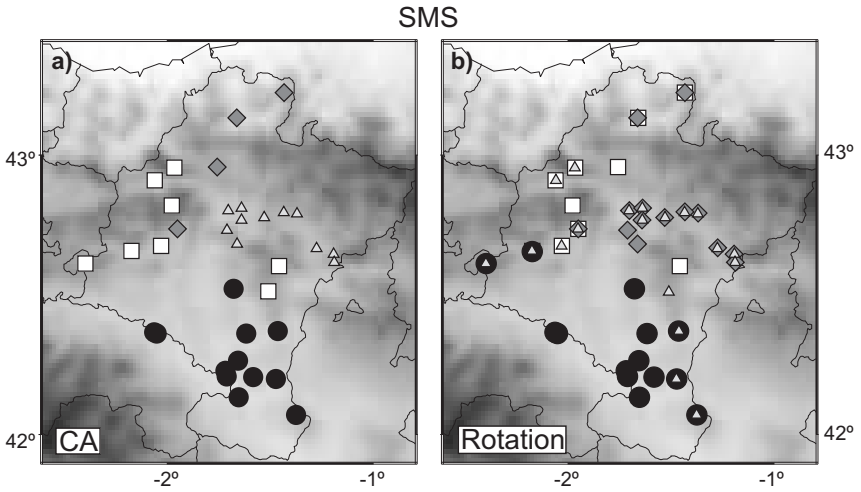

Fig. 5.9: Same as Figure 5.6 but for the SMS dataset (Table 5.1).

regions (Fig. 5.9b) as the RSTMS classification (Fig. 5.6b) and therefore as the one obtained with observations (Fig. 4.8a). These similarities suggest that the identified wind subregions are robust for the whole observational period from 1992 to 2005.

The rotated SMS principal components allow to analyze the temporal wind variability at the subregions over the whole observational period (1992 to 2005) without the handicap of missing data. To illustrate this advantage of the simulation against observations, the rotated principal components calculated with the reduced observational dataset (1999 to 2002) are extended as much as possible and compared against the rotated SMS principal components. With this aim, the standardized anomalies of the observed wind time series for those days with sufficient (more than ten in this case) wind observations available are projected onto the eigenvector defining each observational subregion. The original rotated principal components and the projections are virtually identical in the overlapping parts (1999 - 2002) but the projections cover the complete observational period. The wavelet power spectra of these projected time series and of the SMS rotated principal components are displayed in Figure 5.10. Results for the NS subregion are omitted since the simulation shows problems to reproduce the wind variability at this subgroup (Figs. 5.7c, 5.8c). The wavelet power spectra calculated with observations show several cones of influence as a consequence of the missing data which do not allow to extract conclusions at the lower frequencies over most of the period (left panels in Fig. 5.10). Instead, the power spectra calculated with the simulation do not suffer from this limitation (right panels in Fig. 5.10). Since the rotated principal component associated with the EV is in concordance with both the zonal and meridional wind component of the subregion, the wavelet power spectra of the EV subregion (Fig. 5.10b) shows a similar structure as the one calculated previously with the mean zonal time series at the subregion (Fig. 5.5a). The rotated

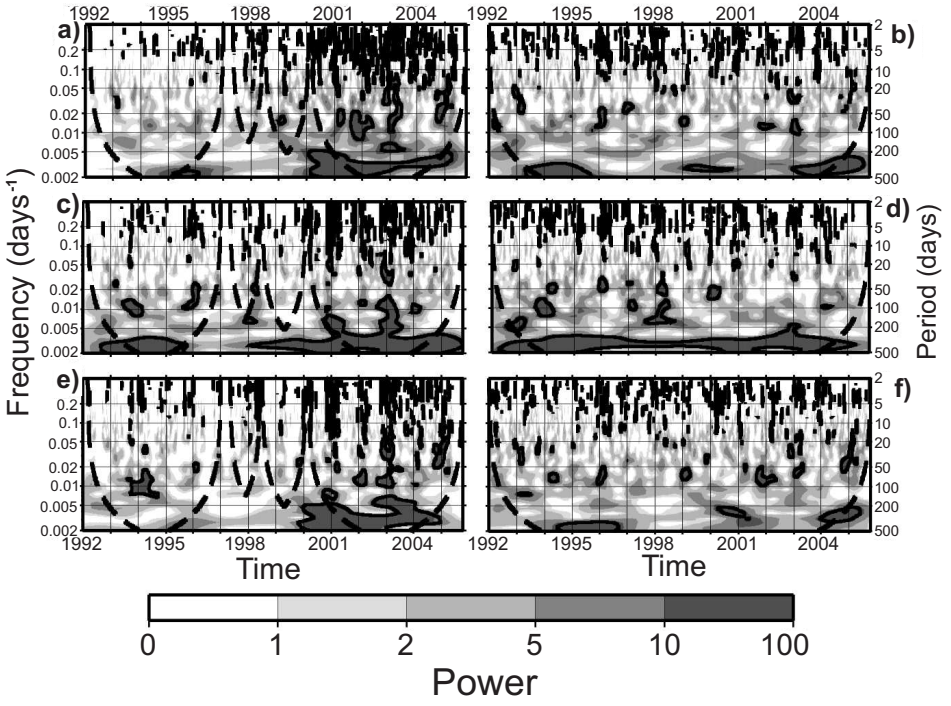

Fig. 5.10: Wavelet spectral power of the standardized projections calculated with the complete observational dataset (left panels) and rotated principal components calculated with the SMS dataset (right panels) for the EV (first row), MS (second row) and NV (third row) subregions. Shading and colours are as in Figure 5.5.

principal component of the MS is in agreement with the zonal component of the subregion and therefore, its wavelet spectra (Fig. 5.10d) is similar to the one calculated with the zonal wind of the subregion (Fig. 5.5b). These similarities provide additional evidence of the realism of the simulation to reproduce the wind variability at these subregions. The similar wind variability at the EV and NV subregions is also captured by the simulation as it is revealed by the similar wavelet spectra of their principal components (Fig. 5.10b,f)

5.4.2 Regionalization over the 1992-2005 period and the full domain

The limited number of observational sites used in the regionalization (Fig. 4.1) leads to questioning the extent to what the spatial sampling affects the final subgroups. Since the numerical simulation shows certain accuracy in identifying the observed subregions and reproducing their temporal variability, one can take benefit of their larger spatial coverage to shed some light on this question. With this aim, the wind regionalization is

reproduced with the complete simulated dataset and therefore avoiding any temporal or spatial mask.

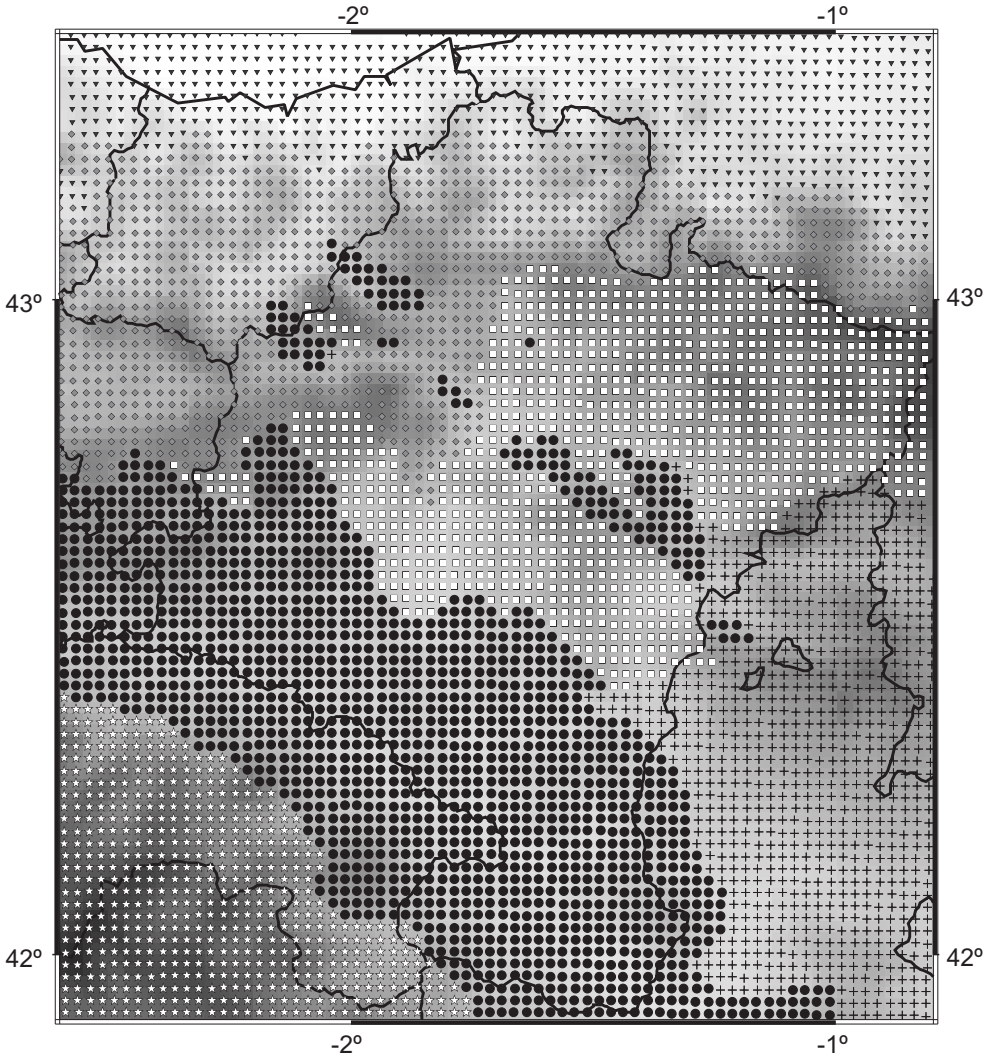

Fig. 5.11: Regionalization configuration attained after applying the methodology based on the CA to the complete simulated dataset (Table 5.1). The simulated grid points with similar temporal wind variability are represented with the same symbol and therefore define the identified subregions.

The regionalization obtained after applying the CA methodology is displayed in Figure 5.11. The resulting subregions show clear resemblances with the regionalization attained with the observed surface wind (Fig. 4.8c) and the SMS (Fig. 5.9a). Additionally, the higher spatial coverage of the simulation allows for the identification of other subregions outside the CFN political boundaries. The EV subregion (circles) is well defined by the Iberian system, the *western* and the *eastern mountains* which shape the basin. The valleys to the north of the Aralar ridge and the NV subregion are also included in this subregion indicating a similar temporal variability of these areas. The similar wind variability of the EV and the NV was already noticed but this time the numerical simulation is unable to identify them as different subgroups. The regionalization identifies a subregion (+ symbol) to the East of the CFN that includes some areas of the Ebro valley where no observations were available. A MS subregion (squares) is identified which includes the Pyrenees, the *eastern mountains* and some regions of the *western mountains*. Another mountain subgroup (stars) located outside the CFN in the Iberian system is also discriminated. A NS subregion is also identified (diamonds), and covers a wider extension than previously detected in the north of the CFN and surrounding areas. Finally, a subregion located to the North of the Pyrenees and Cantabrian mountains outside the CFN political boundaries (triangles) is also isolated.

The wind regions obtained with the alternative regionalization approach based on the PCA-rotation are displayed in Figure 5.12. The subregions identified with this methodology are consistent with those obtained with the CA methodology (Fig. 5.11) thus providing robustness to the results achieved.

5.5 Conclusions

It has been shown that a regional-scale evaluation instead of a site-based evaluation filters out the noise associated with local effects or *representativeness errors* reinforcing the regional signal and therefore providing a more adequate framework to evaluate the simulation performance. The simulation presents better skill in reproducing the regional meridional wind variability than the zonal one. This is a consequence of the similar meridional wind variability observed over the whole study region, attributed to the influence of large mountain systems adequately represented by the simulation. The zonal variability causes the differences between the wind regions and thus, it is more affected by the regional features of the terrain which are not represented as well as the larger scale topographic features. A wavelet analysis showed a reasonably good replication of the observed variability by the simulation. It revealed a permanent annual band for the zonal wind component at the MS subregion, and intermittent ones at the valley subregions and the meridional wind component.

The model performance was also evaluated by applying the regionalization methodologies used to identify the wind regions to a simulated dataset comparable to observations in its temporal and spatial distribution. The regionalization obtained with the simulation identifies the main characteristics of the subregions with some differences in the definition of their boundaries. The simulated variability at the EV, MS and NV subregions is similar to that found in observations. The NS subregion was also identified in the simulation

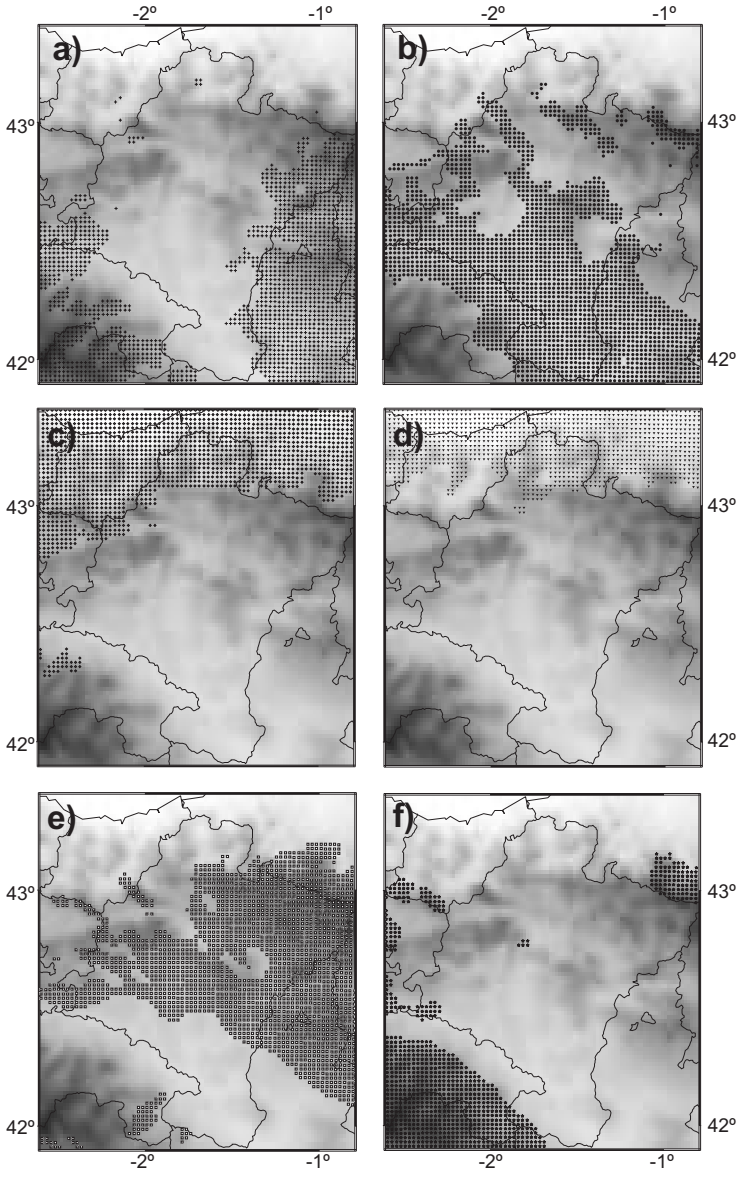

Fig. 5.12: Wind regions obtained after applying the regionalization methodology based on the rotation of the principal modes to the complete simulated dataset (Table 5.1). The symbols represent the distinct subregions.

although with some discrepancies in its temporal variability. This is probably due to the lower number of stations that define this subregion, and the problems of the simulation to reproduce the wind variability at two of its locations, 15 and 26 (Fig. 4.1).

On the basis of the accuracy shown by the numerical simulation to replicate the observed wind variability over the area, the effects of the limited observational sampling of the wind were analyzed. The influence of the observational period length used to identify the subregions (1999-2002) was evaluated by reproducing the classification with the complete simulated time series (1992-2005) at each observational site. This analysis leads to similar subregions, which seem to be robustly defined along the whole complete simulated period. The influence of the limited number of observational sites available to represent the spatial surface wind variability was investigated by reproducing the regionalization with the complete simulated wind field over the CFN and for the complete simulated period. Virtually the same subregions are identified over the CFN as well as additional subregions outside the observational network coverage. This result suggests that the limited sampling of stations seems to reasonably capture the spatial wind variability over the study area.

The results also provide valuable information to improve the observational network. The wind variability at the EV, MS and NV seems to be appropriately captured, but increasing the number of stations at the NS subregion would be desirable. Some improvement of observational density at mountain tops over the Pyrenees to extend the MS coverage would be recommendable. In addition, it would be useful if some information at the subregions identified outside the network coverage is recorded. Since these subregions seem to be well defined covering large areas, the availability of a few stations at each subregion would be enough to capture their wind variability and could provide useful information for future studies.

6

Climatology: wind field classification*

The wind variability over the CFN has been analyzed with observations and with the simulation in Chapters 4 and 5, respectively. This Chapter 6 is devoted to analyze the climatology of the surface circulations. The analysis is performed with the wind observations described in Chapter 2.

The climatological surface flows are analyzed through the identification of WPs that characterize local circulations. A surface wind climatology is provided through some further insight in the main features of the identified WPs such as the magnitude and direction of winds, their annual cycle, frequency, etc. The regional climate variability associated with the different advective properties of the surface circulations are in addition evaluated. This will be accomplished by relating the resulting WPs to other surface meteorological variables, i.e., temperature, humidity, global radiation and precipitation.

The forcing mechanisms that give rise to the already identified wind circulations are analyzed by classifying the synoptic fields governing the low atmosphere circulation and relating them to the WPs. The associations found between the synoptic patterns and the WPs reveals the importance of the ageostrphic balance and the intensification of the surface flows exerted by the pressure gradient along the valleys. In addition, the classifications provide an appropriate framework to analyze the performance of the simulation (Chapter 3) to reproduce the typical surface flows under representative synoptic situations. This will be accomplished in Chapter 7.

6.1 Background

The surface wind circulations over different areas have been classified into WPs to study quite different issues. For instance, in air quality studies it has been illustrated that while some patterns can transport pollutants into a region others bring clear air, thus affecting the concentrations of visibility-reducing aerosols (Green et al., 1992b; Banta et al.,

* The main contents of this chapter are included in:
Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, E. García-Bustamante and J. Navarro, 2009: Climatology of wind patterns in the Northeast of the Iberian Peninsula. *Int. J. Climatol.*, **29**, 501-525.

1999) or toxic air pollutants (Darby, 2005). In the field of renewable energies, knowledge of WPs is particularly important for modeling the efficiency and operation of the energy conversion systems such as solar devices or ventilation designs (Gomez-Muñoz and Porta-Gándara, 2002) and especially for wind energy applications (Palutikof et al., 1987; Burlando et al., 2008). The strongest winds over a region usually present some privileged directions (Zurański and Jaśpińka, 1996) and for this reason, a WP classification can help to understand which flow directions or WPs can potentially produce extreme wind episodes. Knowledge of the regional WPs is also relevant for assessing risks associated to the occurrence of some meteorological situations (Thuillier, 1987). In addition, interactions of wind with the water surfaces are controlled by the WPs which can induce upwelling or downwelling (Bakun and Agostini, 2001; Torres et al., 2003) or can change the evaporation regime (Verburg and Hecky, 2003).

The WP classification of surface wind flows over the CFN is herein performed to analyze the prevailing circulations and their relation to large scale flow. Two different automated approaches are used to identify the WPs. The first methodology performs the classification according to the spatial similarity of the wind fields (Kaufmann and Whiteman, 1999), while the second one is based on their temporal variability (Green et al., 1993). As for the wind regionalization, the use of two methodologies allows for a later evaluation of the consistency of results thus providing confidence to the robustness of the WPs classification obtained.

Quite different approaches have been used to classify synoptic situations (Yarnal, 1993). In particular, several manual classifications of the synoptic fields over the IP have been already performed (Soler, 1977; Linés, 1981; Guardans and Palomino, 1995; Font, 2000). However, manual classifications are based on the experience of the meteorologist and therefore, present a certain degree of subjectivity that can be mitigated employing automated procedures (Yarnal, 1993; Yarnal et al., 2001). Different methods that can help to address this question can be found in the literature. Jenkinson and Collison (1977) developed an automated sea level pressure (SLP) map-pattern classification based on direction and vorticity indices of the geostrophic flow, in an attempt to reproduce previous results from a manual classification over the British islands (Lamb, 1972). The same methodology was applied to other locations (Linderson, 2001) including the IP (Trigo and DaCamara, 2000; Spellman, 2000), but some deficiencies on the methodology have been pointed out for this particular region (Martín-Vide, 2001). Other automated map-pattern classification methods widely employed are the correlation-based methodologies of Lund (1963) and Kirchhofer (1974). These procedures seem to classify the smoother upper air pressure fields more satisfactory than the ones in the low atmosphere, where the present work is focused (El-Kadi and Smithson, 1992). Willmott (1987) recommended the use of alternative classification methodologies like eigenvector techniques. This idea is also supported by Huth (1996), who compared the correlation-based techniques of Lund (1963) and Kirchhofer (1974) against multivariate methods, finding better results with the latter. The eigenvector techniques have been successfully applied in a variety of cases (Davis and Kalkstein, 1990; Kidson, 1994; Esteban et al., 2006) including studies centered in the IP (Zhang et al., 1997; Romero et al., 1999c). Hence, an automated classification based on eigenvector techniques is adopted herein (Romero et al., 1999c)

due to its advantages over the correlation based techniques (Willmott, 1987; Huth, 1996). The SLP and the geopotential height at 850 hPa are the variables involved.

The following sections provide further details on the methodologies selected and the resulting classifications.

6.2 Wind field classification methodologies

In order to accomplish the classification of the daily wind fields into WPs two strategies were explored. The first one groups wind fields according to their spatial similarity and the second one classifies them attending to their temporal variability. Both methodologies employ cluster analysis (CA, Anderberg, 1973) to obtain the WPs, but only the latter involves the use of PCA to isolate the main modes of variation. The use of two methodologies not only allows for a comparison between different techniques, but also helps to discriminate the suitable number of clusters (WPs) in the resulting classification.

To mitigate the possible adverse effects that the missing values could cause during the classification, a reduced dataset is created in a similar way as it was done in Section 4.1 to prepare the observational data for the wind regionalization. A different reduced dataset is here created to take into account the additional wind information available, since the regionalization analysis was accoplished with a previous wind dataset that only had observations until September 2002. The reduced dataset is created with the best 36 stations in terms of the number of available observations (discarding stations 6, 22, 23, 30 and 36). Additionally, only those days with more than 90 % of measurements within the 36 remaining stations were selected (1354 days). The spatial similarity and the temporal variability classification methods are then applied to this reduced dataset obtaining two independent preliminary classifications. Through the comparison of results obtained with both methods, the suitable number of cluster (WPs) is more robustly established and thus the degree of subjectivity reduced. Furthermore, such a comparison allows for the selection of the most appropriate methodology to be employed in the final classification in which each day from the complete dataset is assigned to the most similar WP obtained in the preliminary classification. This implementation of two consecutive classifications was first proposed by Kaufmann and Whiteman (1999) in order to reduce the influence of the missing values.

The final classification of daily wind fields from the complete dataset leans on the WPs obtained from the reduced dataset. Each WP is represented by a class center or centroid defined as the mean field of all the elements in the cluster. The daily fields were compared to each centroid and assigned to a WP accordingly on the basis of a similarity measure. Since the original centroids are defined only over the 36 initial sites, a meaningful representation of data over these 36 stations was required for each daily field in the complete dataset to be assigned to a WP. Thus, only fields with a minimum data availability of 50 % within the reference subset of 36 stations were considered. This decision is rather subjective and is based on previous works (Kaufmann and Whiteman, 1999); it represents a compromise between the desirable classification of all the daily wind fields and the uncertainty associated to the presence of missing data. The measure

of similarity defined to compare class centers and daily wind fields is defined by (Eq. 6.1) and will be introduced in the following sections.

A brief description of the two methodologies used to estimate the preliminary wind field classifications is presented in the following subsections.

6.2.1 Classification based on spatial similarity

This automated classification was developed by Weber and Kaufmann (1995) and improved by Kaufmann and Weber (1996) and Kaufmann and Whiteman (1999). Kastendeuch and Kaufmann (1997) employed this methodology to obtain hourly summer flow patterns in a region over France at different spatial scales. Based on the previous classification of Kaufmann and Weber (1996) over the Alpine region, Weber (1998) obtained a six-year climatology of WPs. A similar study was performed by Weber and Furger (2001) over Switzerland.

Since this approach performs the classification with CA, it requires the definition of a measure of similarity (d_{AB}, distance) between two arbitrary wind fields corresponding to any given days A and B:

$$d_{AB} = \frac{1}{N_{AB}} \sum_{j=1}^{N_{AB}} [(\tilde{u}_{Aj} - \tilde{u}_{Bj})^2 + (\tilde{v}_{Aj} - \tilde{v}_{Bj})^2]^{\frac{1}{2}} \tag{6.1}$$

d_{AB} is the average of the euclidean distances at every site calculated over the N_{AB} zonal ($\tilde{u}$) and meridional ($\tilde{v}$) wind components available at both day A and B. This distance is the same as the one used to identify the wind subregions (Eq. 4.2), but the tilde over the components denotes that in this case they are normalized to reduce potential local and temporal biases on the calculation of distances: first, the components are normalized by the time-averaged wind speed at each site to prevent an over-weighting of particularly windy sites (Weber and Kaufmann, 1995); second, each resulting daily wind component is normalized by the spatial-averaged wind speed over all available stations to ensure that wind fields that differ by an overall scaling factor are grouped together (Kaufmann and Whiteman, 1999).

A two step clustering procedure (Kaufmann and Weber, 1996) similar to the one used to perform the regionalization is employed due to its advantages against the single-step algorithms (see Section 4.2.1). The first step consist on the hierarchical CLA (Johnson, 1967). The second step is a non-hierarchical algorithm similar to the k-means (Anderberg, 1973) used to regionalize, but with a variation consisting in the assignment of fields found to be difficult to classify to a special group of indeterminate wind fields (Kaufmann and Weber, 1996). This is implemented by defining a threshold value, d_{limit}, for the similarity measure (Eq. 6.1). If the distance of a particular wind field exceeds this value to all centroids, that daily field will be assigned to the special group. The elements of this group will not be used to recalculate the centroids in the subsequent reassignment with the k-means algorithm.

The choice of d_{limit} is an heuristic decision which depends on the particular characteristics of the wind datasets. Typical values for d_{limit} are slightly lower than one (Kaufmann and Weber, 1996; Kaufmann and Whiteman, 1999). The value selected here

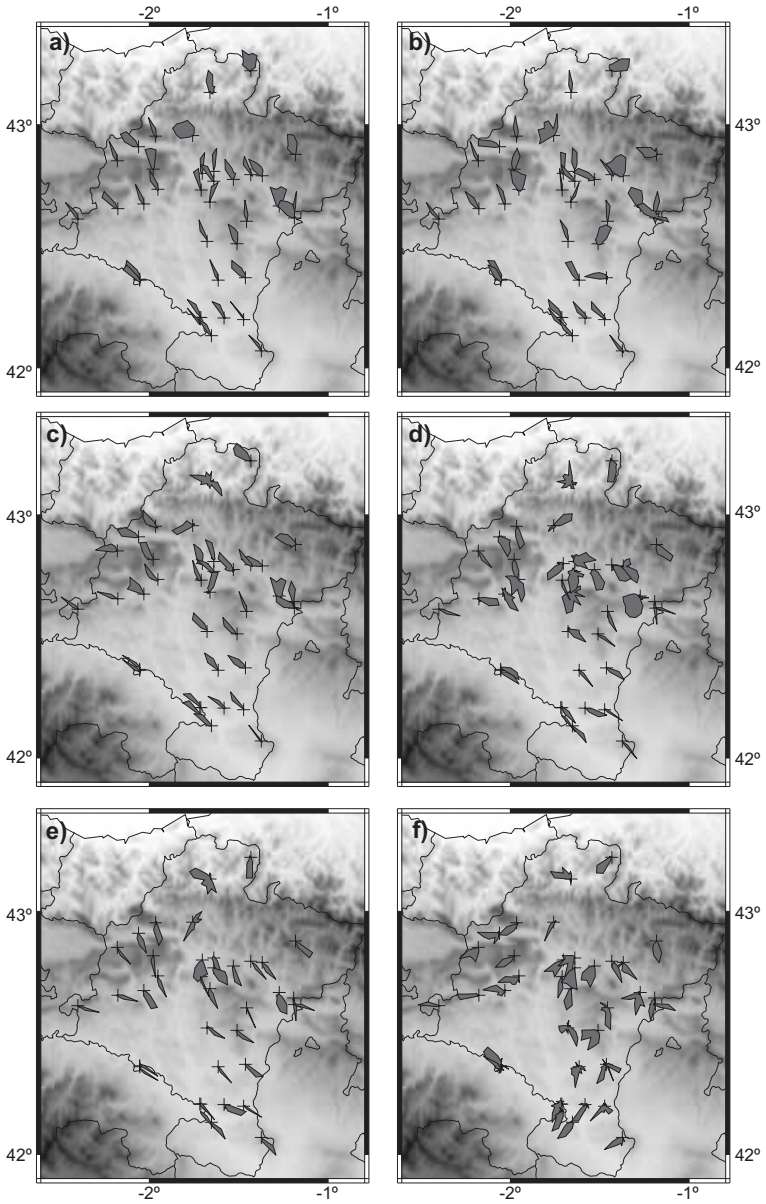

Fig. 6.1: Normalized wind roses of daily wind fields classified as a) WP1, b) WP2, c) WP3, d) WP4 e) WP5 and f) WP6 after applying the methodology based on the spatial similarity to the reduced dataset.

for the present classification after an exploratory analysis is $d_{limit} = 0.7$. Additional comments concerning this decision will be made in Section 6.3.

6.2.2 Classification based on temporal variability

The second automated strategy classifies wind fields into WPs according to their temporal variance. A vectorial PCA is applied to the correlation matrix as a first step (T-mode; see Richman, 1986). In this decomposition, the principal components (scores) contain the spatial variation of the temporal eigenvectors (Green et al., 1993). Hence, the typical CA of the eigenvector components (loads) yields groups with similar spatial patterns (Romero et al., 1999a). Such an approach can be considered equivalent to a CA applied to S-mode loads used to classify the locations with similar temporal variability in the regionalization (Section 4.2.1).

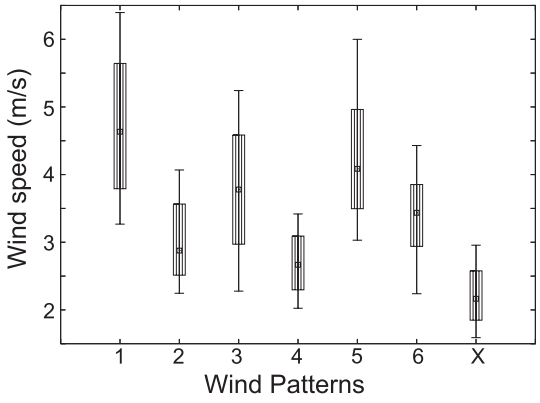

Fig. 6.2: Box and whiskers plots showing the distribution of the mean wind speed field for each WP after applying the classification based on spatial similarity to the reduced dataset. The dot indicates the median, the lower bases of the boxes are the 25 percentiles and the upper ones the 75 percentiles. The lower and upper whiskers indicate the 10 (lower) and 90 (upper) percentiles. The X denotes the special group of indeterminate wind fields.

The CA of the loads is performed as in the previous methodology. In a first step the CLA finds the appropriate number of groups and their initial state values. In a second step the modified k-means method that allows for the assignment of dissimilar days to a special group, reorders the daily fields. However, a slight modification with respect to the previous classification needs to be introduced. Since the eigenvectors contain the temporal information, their loads are scalars and the similarity measure defined in Eq. 4.2 cannot be employed. Therefore, another definition of similarity must be selected. This is overcome by replacing the previous distance (Eq. 6.1) with the Euclidean distance.

Changing the measure of similarity involves a change in the parameter d_{limit} employed by the modified k-means algorithm. In order to produce comparable classifications, d_{limit} is fixed in such a way that a similar number of days than in the previous classification are assigned to the indeterminate group.

6.3 Identification of wind patterns

The outcomes from both the spatial similarity and temporal variability methodologies are compared in order to define the more appropriate number of WPs. The suitable number appears to be six patterns for the former method and five for the latter. In the following subsections, results from the preliminary and final wind fields classifications are presented and discussed.

6.3.1 Preliminary classification based on spatial similarity

The wind roses of the six WPs found by means of the spatial similarity classification method are displayed in Figure 6.1 and the distributions of the mean wind speed fields corresponding to each pattern appear in Figure 6.2. The strong influence of orography is evidenced, channeling surface circulations around the *eastern mountains* and along the Ebro valley as well as along the valleys in the South of the *northern mountains*. The first three patterns (i. e., WP1, WP2 and WP3 Fig. 6.1a,b,c) present northwestern circulations. WP1 and WP2 share some spatial characteristics but several wind roses show differences associated to their orientation and width. The wider wind roses in WP2 (Fig. 6.1) could be related to the weaker forcing of moderate winds (Fig. 6.2) that produce a less defined prevailing wind direction. Further details of this feature will be discussed through its relation with the synoptic circulation in Section 6.6. WP3 also presents northwestern flows. This pattern would have appeared in combination with WP1 in the event that only five patterns would have been allowed to form. In spite of the similarity between both patterns, WP3 shows a more zonal orientation of the wind roses than WP1. This is especially visible at the northwestern stations. WP4 and WP5 present southeastern flows (Fig. 6.1d,e). The main difference between them is that wind roses are wider in WP4, particularly at the center of the region, associated with its weaker wind speed (Fig. 6.2). Finally, WP6 presents southwestern flows over most of the CFN (Fig. 6.1f).

18.8 % of the 1354 classificable days that constitute the reduced dataset are assigned to the special group of indeterminate wind fields. These daily fields present weak wind speeds over the region (Fig. 6.2) and thus, the wind direction is less precisely defined leading to large distance values with the rest of wind fields. These weak daily wind fields can be associated with relatively calm days or transitions between the WPs. If more fields were classified in an attempt of reducing the number of days belonging to the indeterminate group (increasing d_{limit}) it would result in an undesirable degradation of the wind roses (not shown). This degradation especially affects WP4 and WP6 where several wind roses show various main directions. On the other hand, increasing the number of classificable wind fields by increasing d_{limit} could potentially produce new WPs. However, when that was done, the resulting patterns were essentially the same as reported above with a similar degradation in the wind roses (not shown). Nevertheless, the 18.8 % of days temporally located in the undeterminated group will be reassigned to the other WPs in the final classification.

6.3.2 Preliminary classification based on temporal variability

The second classification procedure applies PCA to calculate the main modes of variation and then performs the CA classification. The explained variance of the leading PCA modes is shown in Figure 6.3. The first mode explains more than 62 % of the variance with a large decrease for the subsequent modes. There is a break on the slope which suggests four as a reasonable number of modes to retain. Therefore, the four principal main modes which take account for 87 % of the total variance were employed in the CA.

Five WPs reveal qualitatively the best agreement with the previous classification (Fig. 6.4). The three northwestern WPs in that case (Fig. 6.1a,b,c) are in quite good agreement with the first three new WPs (Fig. 6.4a,b,c). In contrast, there is only one single pattern with southeastern flows in the present case (Fig. 6.4d). This pattern resembles WP4 and WP5 obtained with the previous methodology (Fig. 6.1d,e). Finally, the last WP (Fig. 6.4e) shows similarities with the previous WP6 (Fig. 6.1f) though some wind roses present somewhat more degraded main directions. The distributions of the mean wind speed fields for each WP are in concordance with those obtained in the previous classification (Fig. 6.2), although there is less wind variability among the WPs (not shown).

As in the previous methodology, the attempt of classifying more wind fields produces a degradation in the wind roses. If six WPs were generated to reproduce the previous results, another pattern with northwestern circulations very similar to the already existing ones would appear.

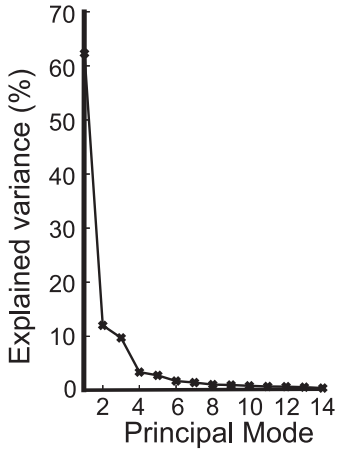

Fig. 6.3: Explained variance of the leading PCA modes of the wind data.

6.3.3 Final wind field classification

Both methodologies produce similar WPs when they are applied to the reduced dataset (Figs. 6.1 and 6.4). This suggests that the identified WPs are robust, and, on the basis of this information the selection of an specific classification is somewhat arbitrary. The main difference between them is the isolation of an additional WP showing southeastern circulation (Figs. 6.1d,e) within the approach based on spatial similarity. This WP does not emerge with the classification based on temporal variability which shows just a single WP with southeastern flows (Fig. 6.4d). The two southeastern patterns identified by the spatial analysis (WP4 and WP5) present different intensities in the surface circulations (Fig. 6.2). Since it seems reasonable to employ the classification which provides more information concerning the surface flows, the classes generated with the spatial similarity approach are those selected for the final classification. It will be illustrated further on that the southeastern WPs present differences in their atmospheric states as well as a distinct

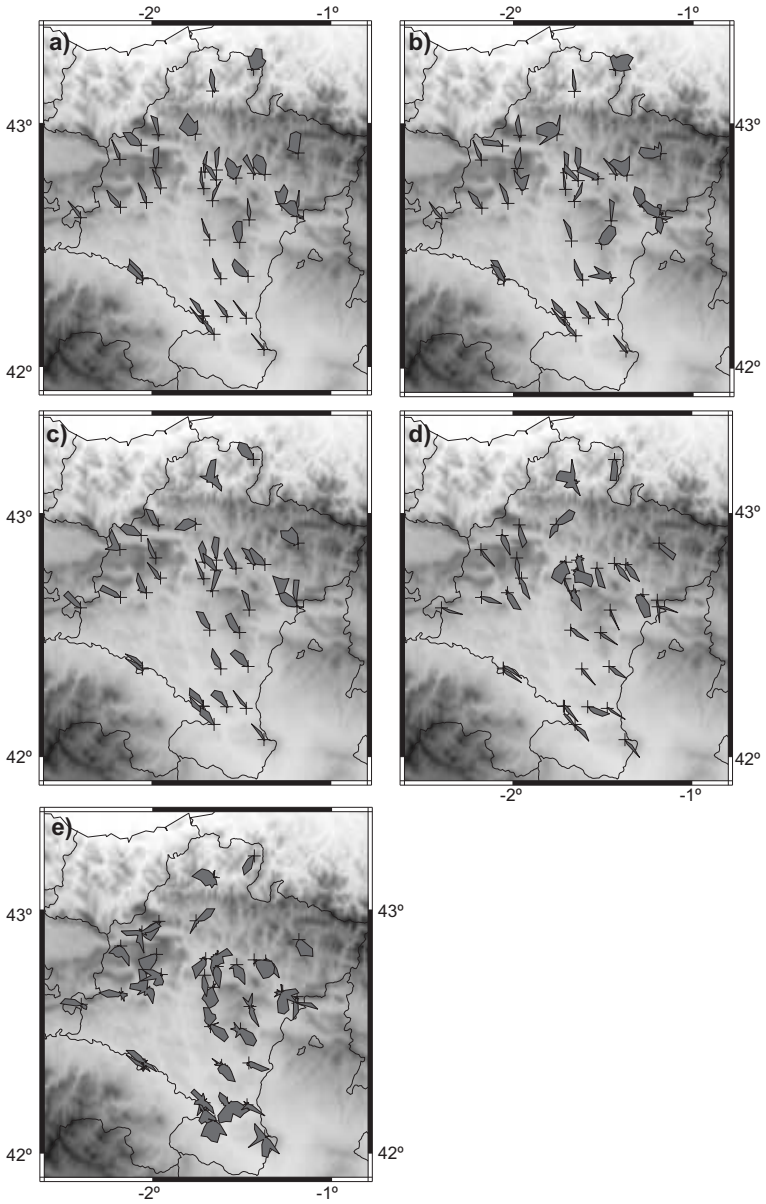

Fig. 6.4: Normalized wind roses of daily wind fields classified as a) WP1 b) WP2 c) WP3, d) WP4 and e) WP5 after applying the methodology based on the temporal variability to the reduced dataset.

appearance in their frequency distribution along the year associated with the main forcing synoptic patterns; these dissimilarities support their treatment as independent patterns. Figure 6.5 shows the mean normalized wind vectors (centroids) corresponding to each WP in the selected classification (Fig. 6.1). Some additional features about local and regional flows in addition to the mean directions presented in Figure 6.1 are obtained that also allow for understanding the differences between both southeastern circulation types. For instance, panel Figure 6.5d highlights lower wind speed averages for WP4 in the *western mountains* and north of the *eastern mountains* than in WP5 (Fig. 6.5e). WP6 (Fig. 6.5f) depicts lower wind speeds in the valley than in the mountain regions.

For a more realistic representation of the mean behavior corresponding to each WP, Figure 6.6 shows the average wind vectors corresponding to all daily wind fields in each WP. Since this patterns are not normalized this allows for considering the diverse ranges of variability in the region. The circulation along the Ebro valley is enhanced in WP1 and WP3 as well as the winds from the southeast in WP5. Some of these features like the higher winds in WP5 compared to WP4 can be easily related to the wind speed distribution for each WP in Figure 6.2.

The resulting wind roses of the six WPs after assigning all the wind fields within the full dataset (see section 6.2 for methodological aspects) are displayed in Figure 6.7. A total of 4050 days (approximately 11 years) are classified. The final appearance of the WPs is in quite good agreement with that of the patterns employed as centroids (Fig. 6.1). However, WP4 and specially WP6 show less unidirectional wind roses. The degradation resembles that of the exploratory classification of wind fields from the indeterminate group (not shown). This fact suggests that the degradation observed in the wind roses is more related to the wind fields dissimilarities than to the presence of missing values in the full dataset. The stations 6, 23, 30 and 36 discarded in the reduced dataset, present wind roses in concordance with the main direction of the WPs flow. Only the discarded station 22 (Fig. 1.2) seems to present a decoupling from the main regional flow probably due to a local channeling.

Some characteristics of the final WPs are displayed in Table 6.1. There main flow direction (NW-SE) accumulates a frequency of 60.9 % in the northwestern sense (WP1, WP2 and WP3) and 30.5 % in the southeastern one (WP4 and WP5). This corresponds well with the main direction of the mountain ridges surrounding the CFN (Fig. 1.2), this fact shows the strong influence of the orography in channeling the general winds. The only WP with a circulation different from the NW-SE direction is WP6 with southwestern flows and a frequency of 8.5 %.

WP1 is the most frequent and the one with the highest wind velocity (Table 6.1). The other northwestern patterns (WP2 and WP3) show smaller wind speeds, especially WP2, where the mean wind speed is almost 2 ms^{-1} weaker than in WP1. In the case of the southeastern patterns, WP5 presents considerably higher wind speeds than WP4. WP6, with a southwestern direction, shows moderate winds. The standard deviation shows a direct relationship with the mean in concordance with the previous analysis (see discussion of Fig. 4.4).

In order to illustrate the atmosphere state related to each WP, anomalies of the wind vector, wind speed, temperature, relative humidity, global radiation and precipitation were calculated and the mean anomaly field for days belonging to each WP displayed in

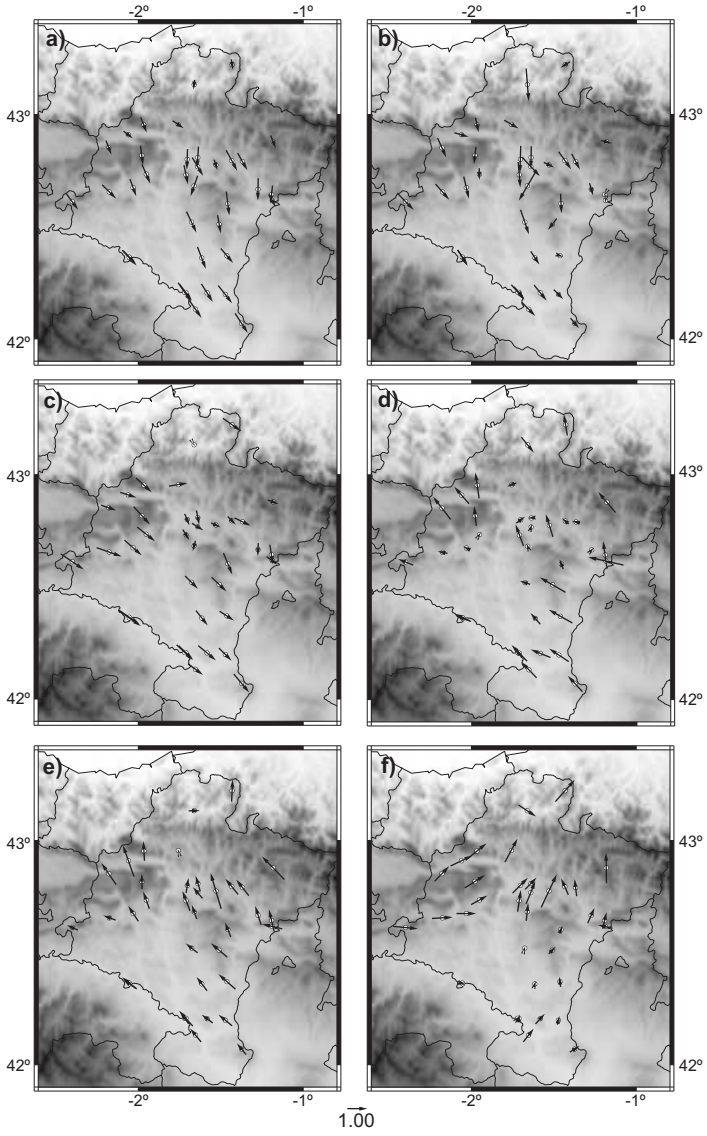

Fig. 6.5: Mean normalized wind vectors (centroids) of daily wind fields classified as a) WP1 b) WP2 c) WP3, d) WP4, e) WP5 and f) WP6 after applying the methodology based on the spatial similarity to the reduced dataset.

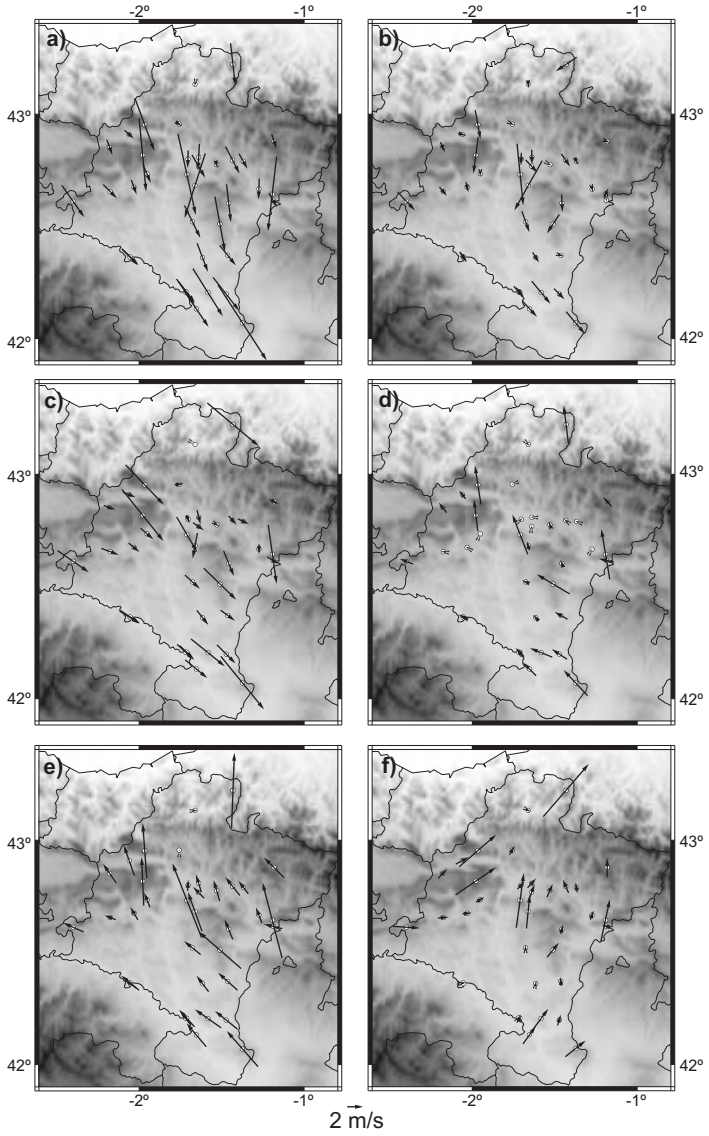

Fig. 6.6: Mean wind vectors of daily wind fields classified as a) WP1 b) WP2 c) WP3, d) WP4, e) WP5 and f) WP6 after applying the methodology based on the spatial similarity to the reduced dataset.

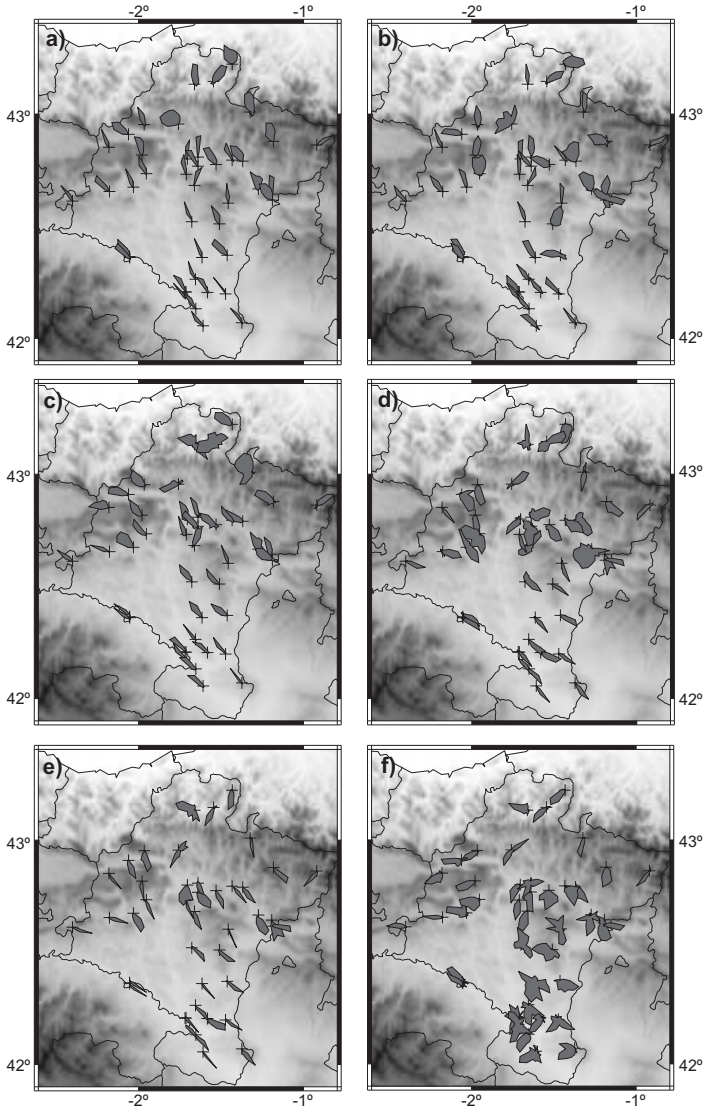

Fig. 6.7: Normalized wind roses of daily wind fields classified as a) WP1, b) WP2, c) WP3, d) WP4, e) WP5 and f) WP6 using days from the full dataset and the spatial similarity based WPs as centroids (Fig. 6.6).

Table 6.1: Final WP's relative frequency, main direction of the flow, mean and standard deviation of wind speed distributions (Fig. 6.7).

Wind pattern	Frequency (%)	Main direction	Mean (m/s)	Std. (m/s)
WP1	25.4	NW	5.0	1.3
WP2	19.3	NW	3.1	0.7
WP3	16.2	NW	3.8	1.2
WP4	15.7	SE	2.7	0.5
WP5	14.8	SE	4.7	1.2
WP6	8.5	SW	3.2	0.9

Figures 6.8-6.13. The same analysis is undertaken with the data from the ECMWF to compare them with observations and is also displayed in Figures 6.8-6.13. Anomalies were computed by subtracting the corresponding long term monthly mean from the daily data within each month in order to suppress the deterministic annual cycle and to highlight relative deviations respect to the mean climate. The mean anomaly wind vectors of the final WPs are in agreement with the ECMWF data which provides an overview of the flow over the IP (Figs. 6.8-6.13a). In general, the advective properties of the WPs are different and in most of the cases are in good concordance with the analysis data. The specific advective properties of each particular WP and the ability of the ECMWF analysis data for reproducing them will be discussed in Section 6.6 and related to their forcing synoptic patterns.

The monthly frequency of every WP is represented in Figure 6.14. Each WP tends to appear preferably at specific months, thus revealing the presence of an annual cycle. WP1 is present almost the whole year, but predominantly from February to August. WP2 discloses its maximum frequency of occurrence in Summer and the minimum during the Winter, being the opposite true for WP3. WP4 though more frequent from May to October appears during the whole year. Finally, WP5 and WP6 tend to emerge during the Winter rather than the Summer. This reveals a seasonality in the surface circulations which ultimately is attributed to the synoptic scale variability which also will be addressed in the next subsection.

6.4 Synoptic classification methodology

For the map-pattern synoptic classification an eigenvector-based methodology is used. The procedure is similar to the wind field classification based on the temporal variability though in this case the scalar PCA is employed. This kind of methodology has been successfully applied in different studies (Davis and Kalkstein, 1990; Kidson, 1994; Zhang et al., 1997; Romero et al., 1999c; Esteban et al., 2006). The map-pattern classification employed here consists in a PCA followed by a CA (Romero et al., 1999c). In a first step, a PCA is applied to the correlation matrix (T-mode) of the selected synoptic variable; then, the most relevant principal modes are retained with the scree test of Cattell (1966) and finally, a two step CA methodology is applied to their loads to generate the patterns.

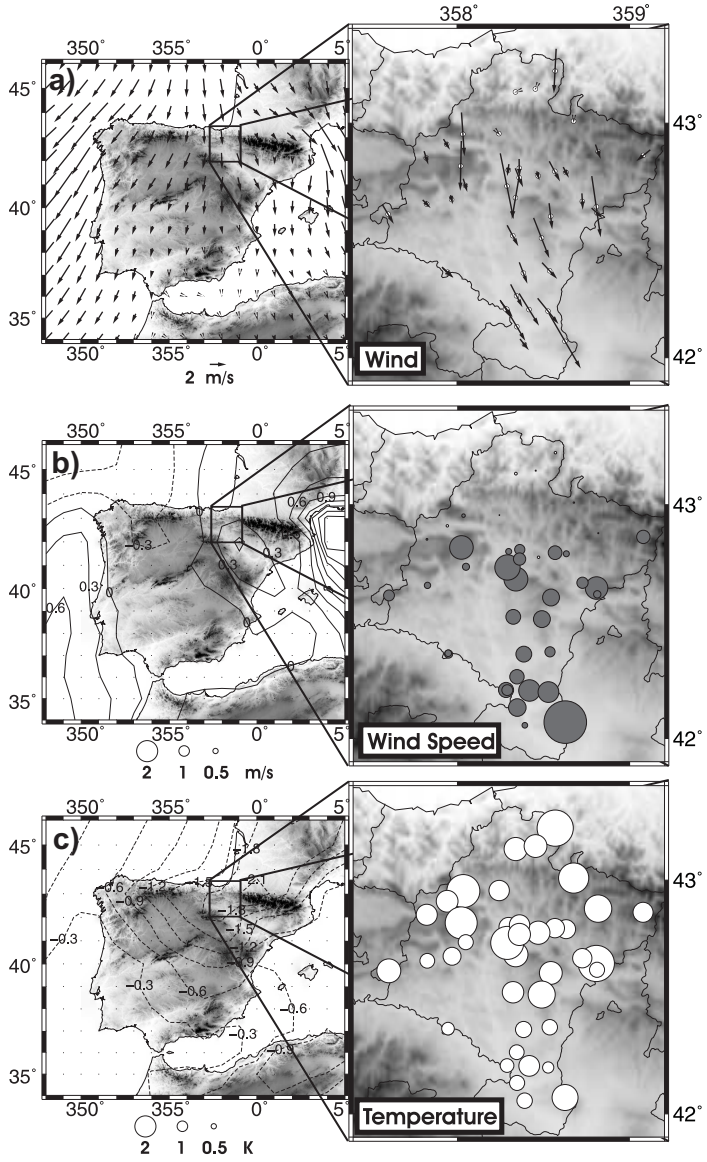

Fig. 6.8: Mean anomalies of the wind vectors (a), wind speed (b), temperature (c), relative humidity (d) global radiation (e) and precipitation (f) of days classified as WP1. Left panels for the ECMWF reanalysis/analysis data with solid (dashed) contour lines representing positive (negative) anomalies. Right panels for the observational data (circles). The radius of the circles is proportional to the magnitude of the anomalies whereas the color indicates the sign (dark for positive anomalies and white for negatives).

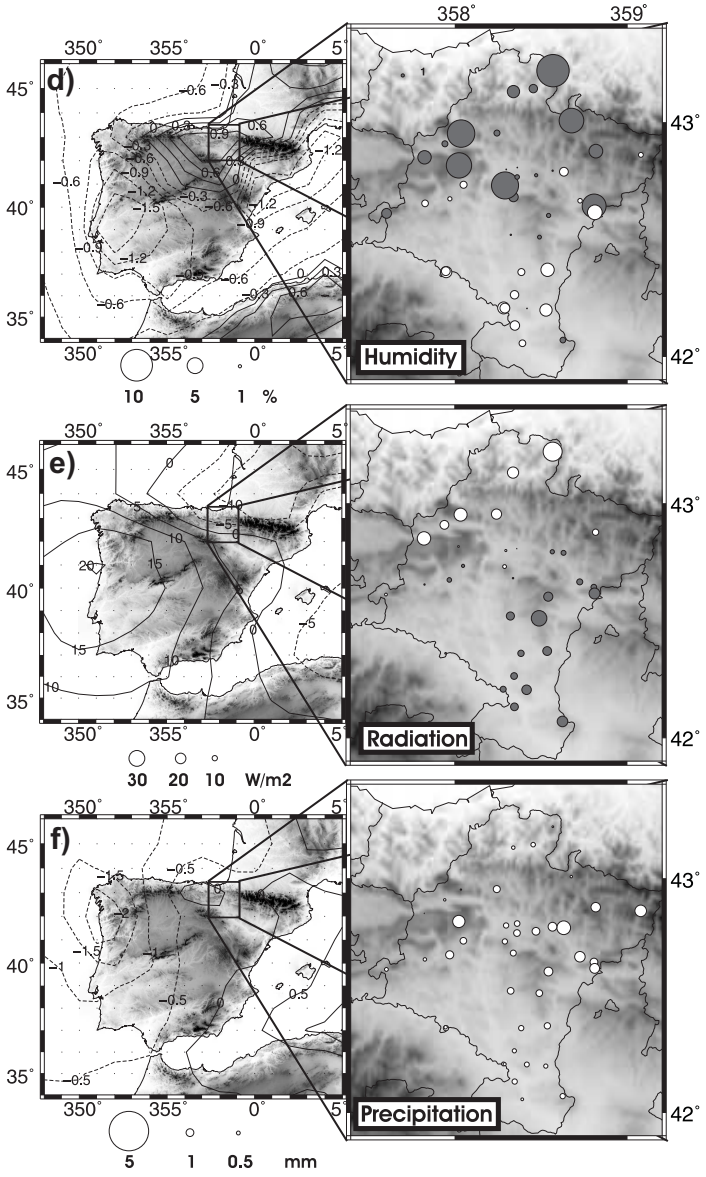

Fig. 6.8: Continuation.

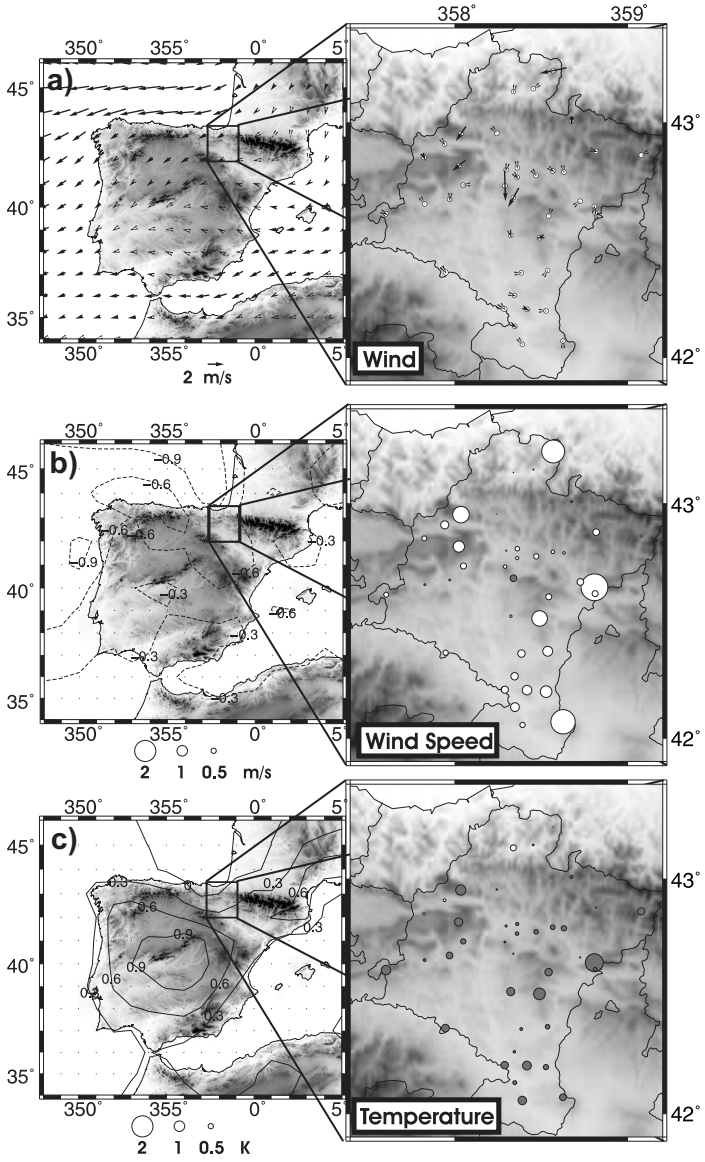

Fig. 6.9: Same as Figure 6.8 but for days classified as WP2.

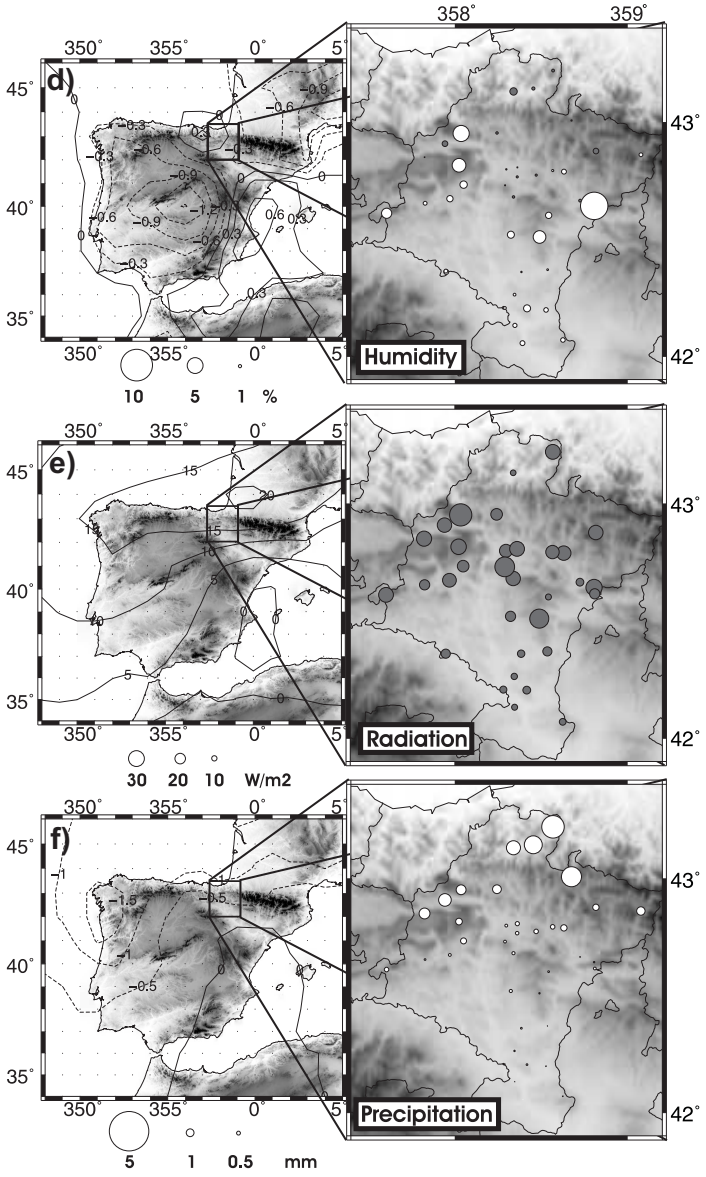

Fig. 6.9: Continuation.

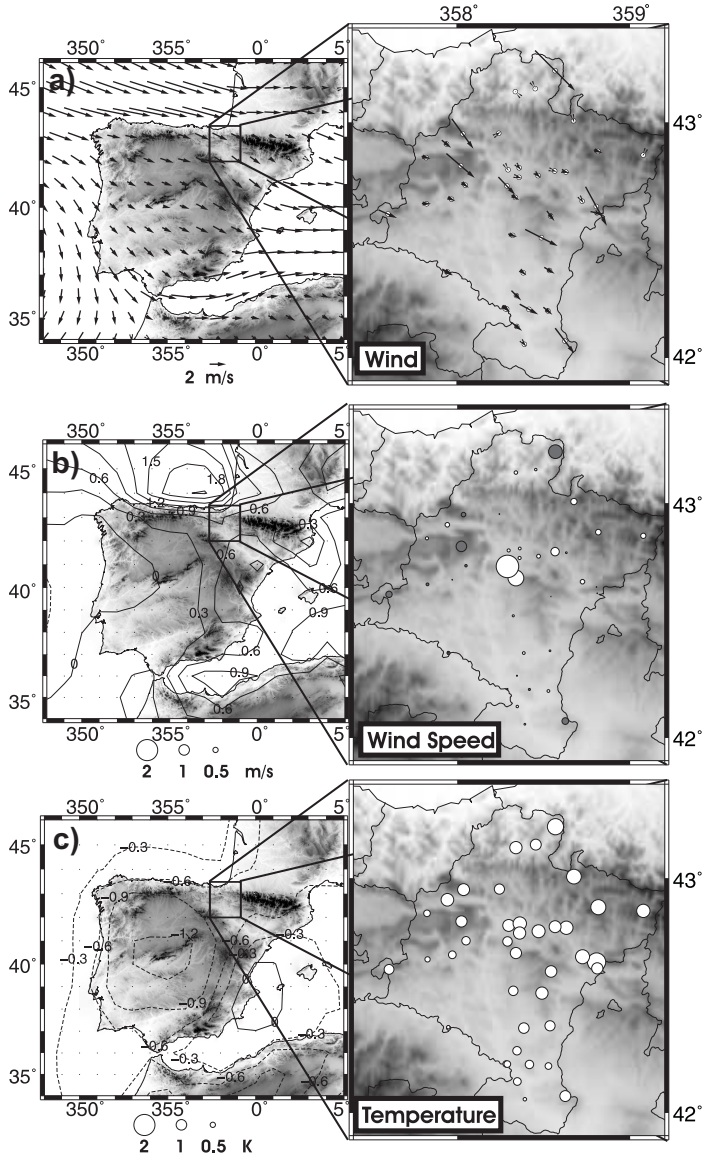

Fig. 6.10: Same as Figure 6.8 but for days classified as WP3.

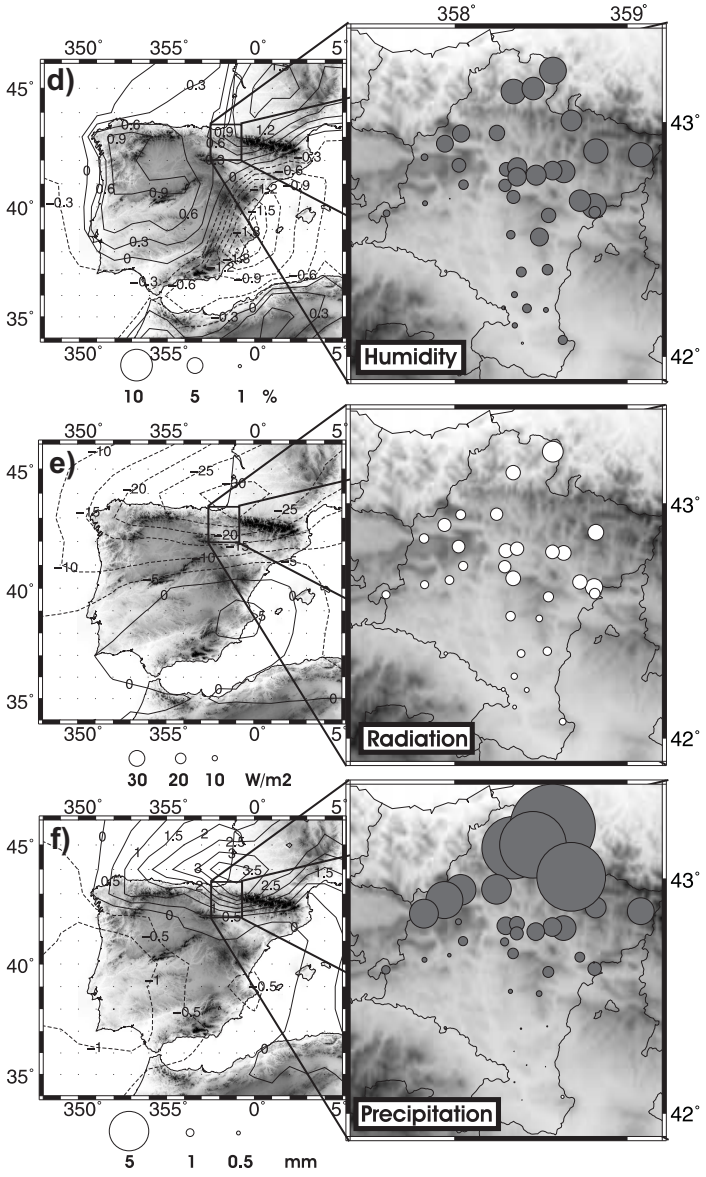

Fig. 6.10: Continuation.

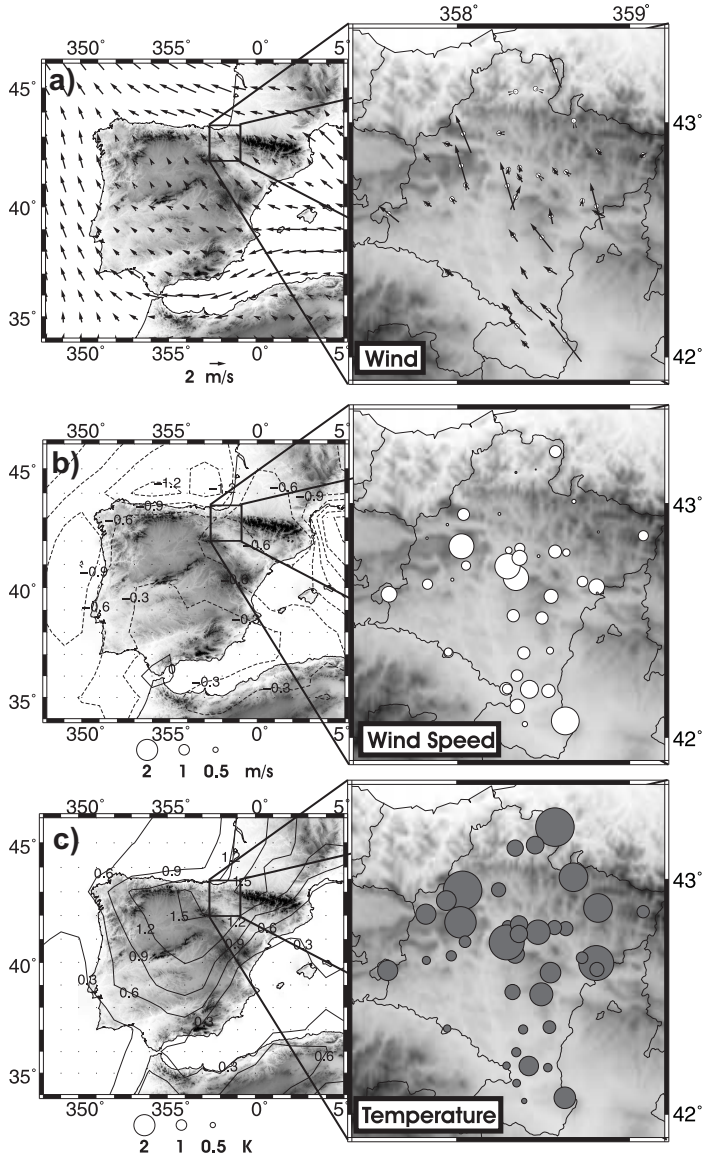

Fig. 6.11: Same as Figure 6.8 but for days classified as WP4.

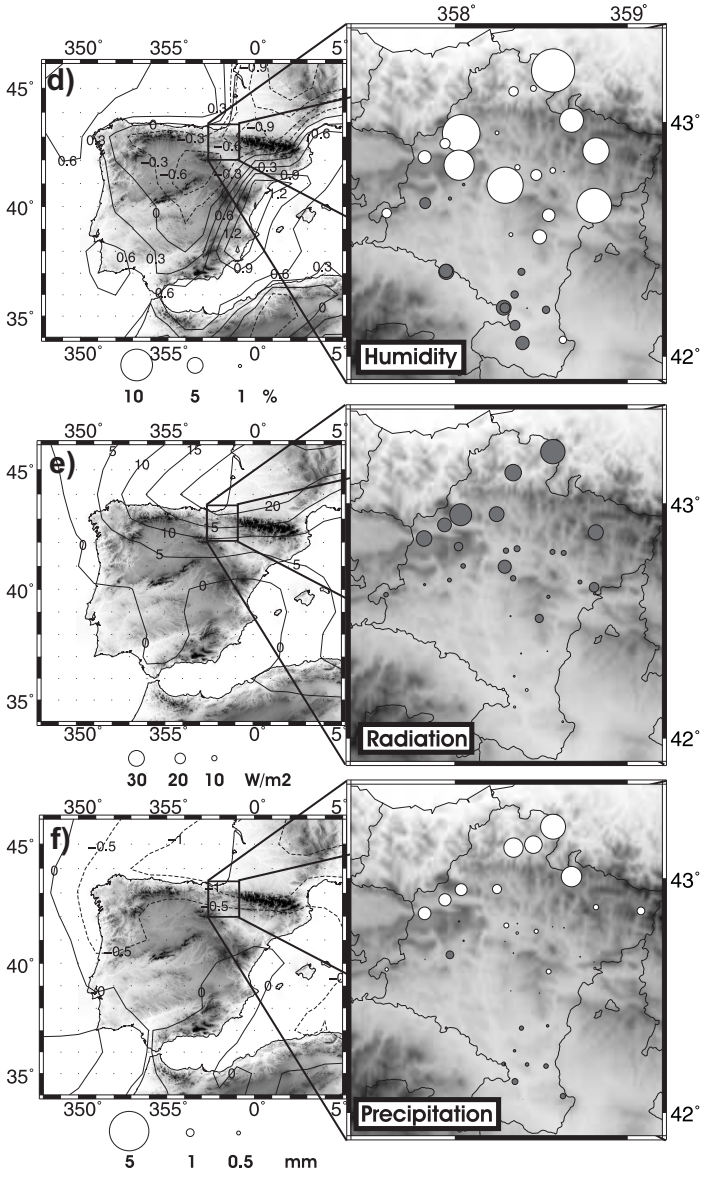

Fig. 6.11: Continuation.

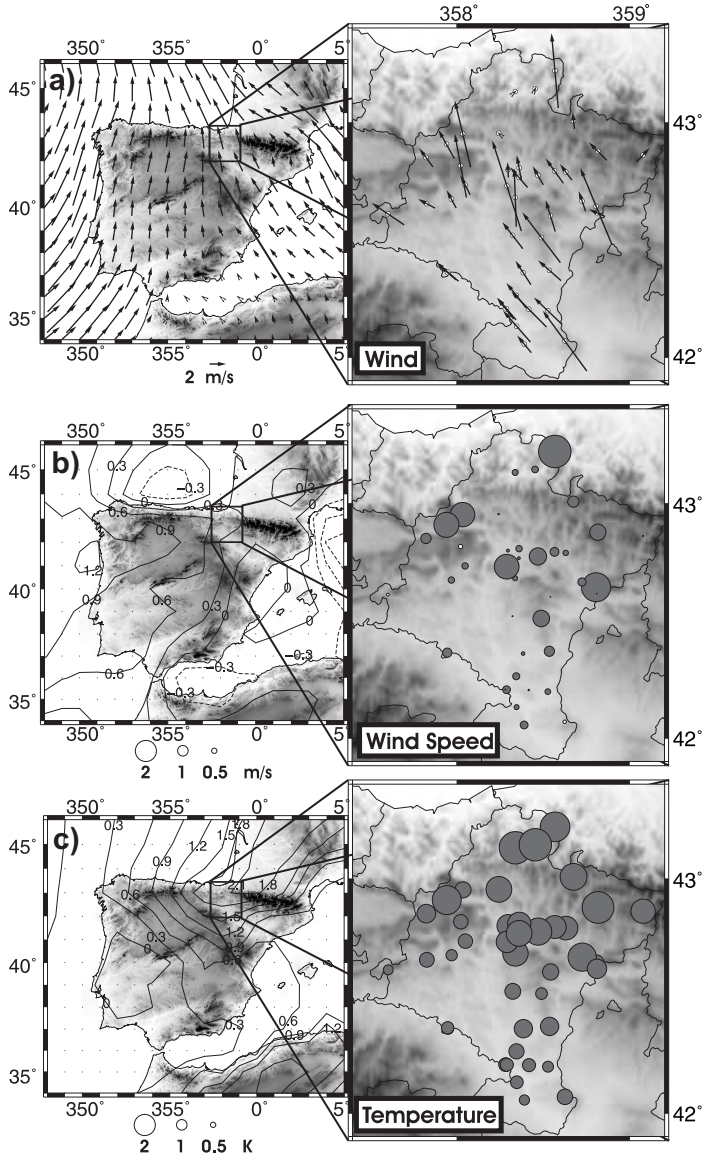

Fig. 6.12: Same as Figure 6.8 but for days classified as WP5.

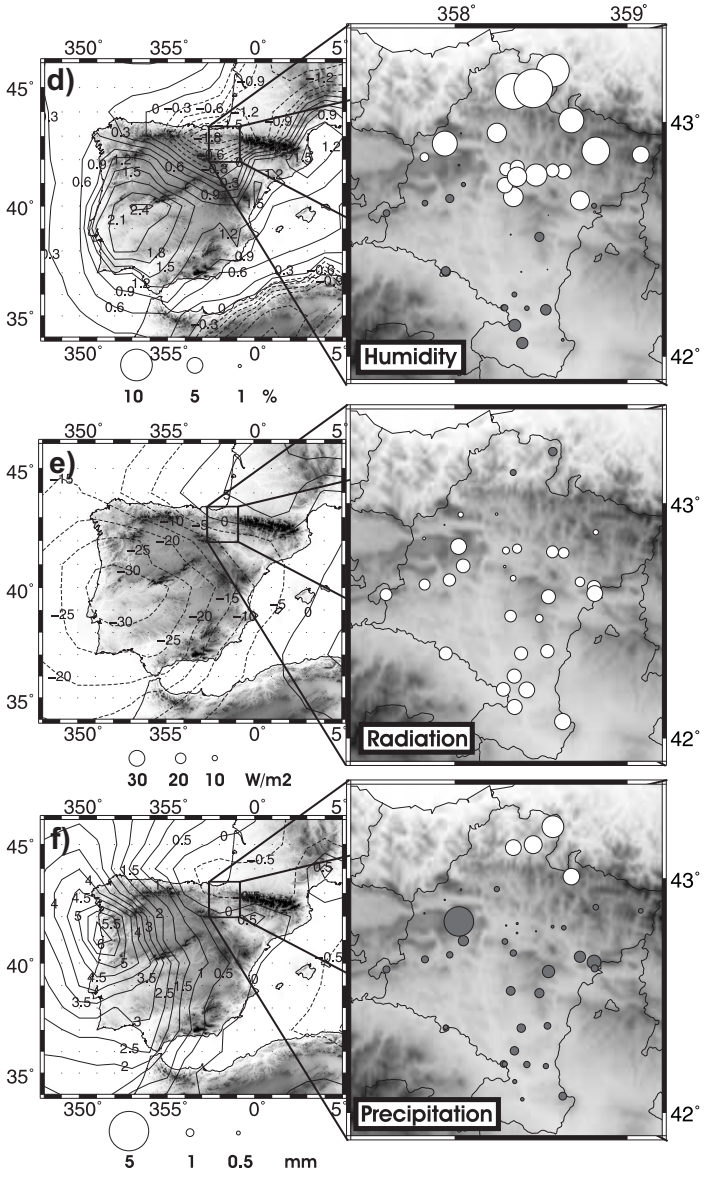

Fig. 6.12: Continuation.

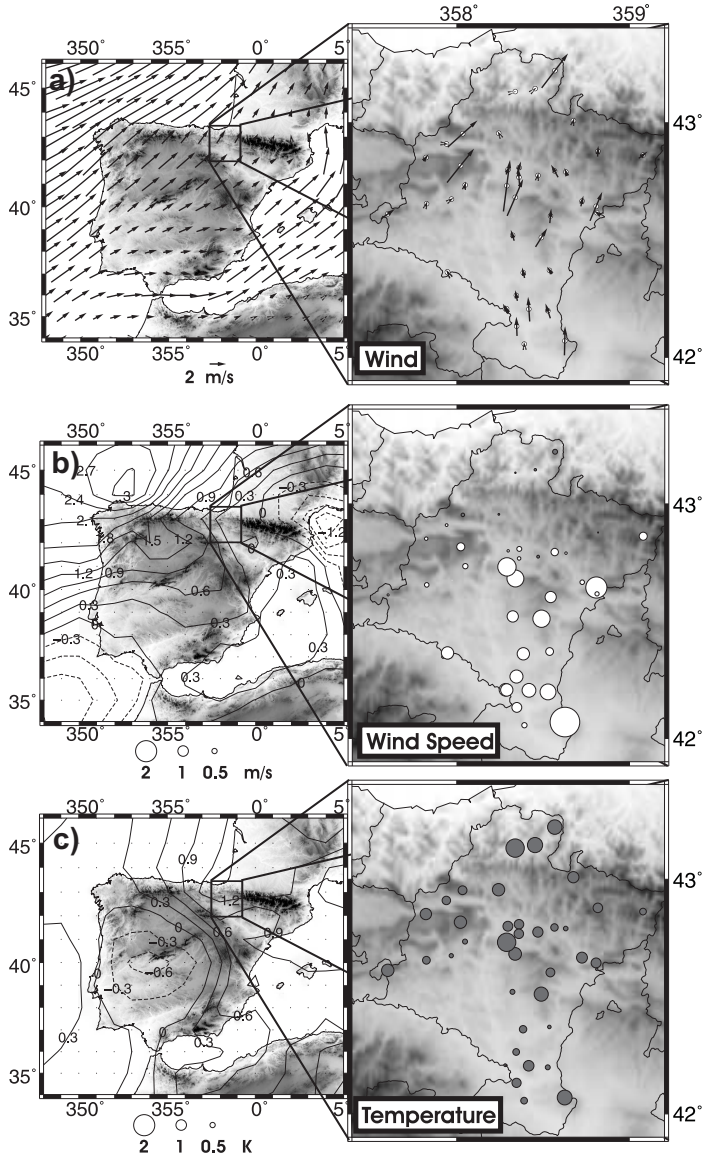

Fig. 6.13: Same as Figure 6.8 but for days classified as WP6.

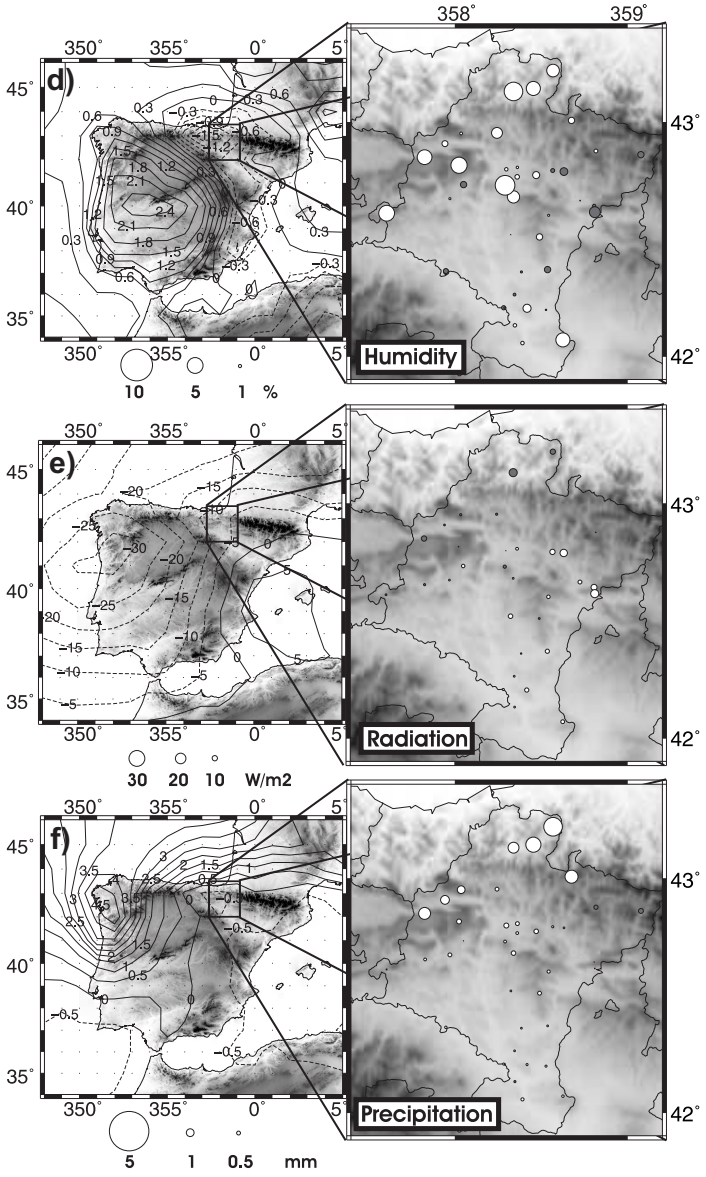

Fig. 6.13: Continuation.

Since the loads are scalars, the Euclidean distance is again used as the similarity measure. This allows for the adoption of the Ward's algorithm (Ward, 1963), that is appropriate to find initial state values, for the first step of the CA classification. The second step of the CA is performed with the k-means algorithm completing the methodology described for instance in Romero et al. (1999c).

Therefore, the main difference between this classification and the wind field classification based on the temporal variability is the hierarchical algorithm employed in the first step of the CA, Ward's algorithm in the synoptic classification and Complete Linkage in the wind field classification. Although technically it would be possible to employ the Ward's algorithm in the classification based on the temporal variability, the Complete Linkage was selected to be in concordance with the wind field classification based on the spatial similarity wherein the Ward's algorithm could not be employed (Section 6.2).

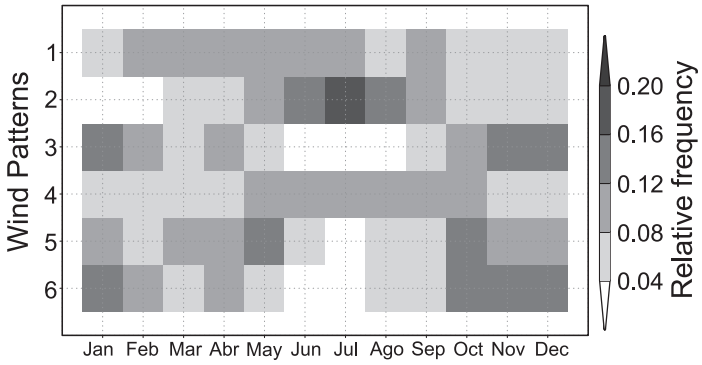

Fig. 6.14: Monthly frequency of the six WPs shown in Figure 6.7.

Since the purpose of the analysis is to relate the resulting synoptic patterns to the surface WPs, synoptic variables governing low atmosphere circulations like SLP and geopotential height at 850 hPa, were tested using several geographical windows with different sizes.

At this point it is worth commenting the two approaches that can be taken to analyze the forcing mechanism of surface variables. They are referred by Yarnal (1993) as the *circulation-to-environment* classification and the *environment-to-circulation* classification. In the first one, a classification of the synoptic circulations is first performed and then related to the surface characteristics of the selected variable (the surface circulations in the present case). In the second approach, the classification of the surface properties is first performed and then the synoptic situation of the different surface patterns evaluated. Both approaches can be found in the literature in relation with a classification of surface circulations into WPs and the analysis of their forcing synoptic patterns. For example Weber (1998) applied the *circulation-to-environment* approach and Pandžić and Likso (2005) the *environment-to-circulation* method. In the first approach the synoptic classification is independent of the environmental response whereas in the second case

the synoptic classes are conditioned to the various regional types. The *circulation-to-environment* approach is here adopted performing the independent synoptic classification as described in the previous paragraphs. This decision was partially based on the benefits that an independent classification of the synoptic circulations provide in the causality of the emergent associations with the surface patterns. In addition, the selection of the *circulation-to-environment* approach was taken on the basis of some previous works that did not find a unique relationship between the synoptic classes and the surface flow patterns such as the mentioned work of Weber (1998) or the study of Conil and Hall (2006). This suggests that the *environment-to-circulation* approach would be somewhat misleading since the averaged synoptic fields associated with each WP could be a combination of quite different structures. This can happen when regional features dominate over the large scale flow. For instance, we can consider the effects that orography produces over the large scale circulations by modifying and channeling them and therefore reducing the degrees of freedom of the surface flows. In fact, the latter is magnified in complex terrain regions as it is the case of the CFN (Weber and Kaufmann, 1995).

6.5 Identification of pressure patterns

The SLP fields over the IP were classified and related to the surface circulations over the CFN in order to understand their forcing mechanisms. 850 geopotential hPa fields were also considered and subjected to classification. However, the most unequivocal relationships between the mentioned synoptic fields and the surface circulations (WPs) were found for the SLP and therefore only results for this variable are shown from here on.

The influence on the synoptic classification due to the size of the geographical window was also evaluated. The smallest window tested, which covers the IP as in the studies performed by Romero et al. (1999c) and Spellman (2000) to obtain their synoptic classifications, produces better results. The rest of the geographical extensions considered cover larger areas and therefore are affected by some large scale circulation features that impact the classification degradating the relation between the pressure patterns (PPs) and the regional WPs. Such an effect was also noticed by Conil and Hall (2006).

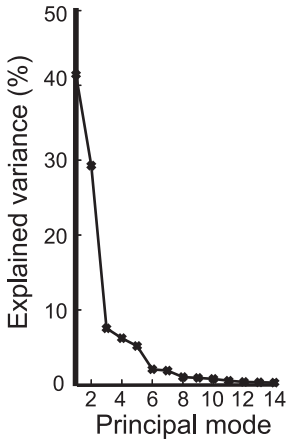

Fig. 6.15: Explained variance of the SLP leading PCA modes.

The explained variance of the leading modes after applying PCA to the SLP correlation matrix is displayed in Figure 6.15. The slope breaks on at the number of six which suggests this is a reasonable number of modes to retain for the CA classification.

The emergent PPs after performing the CA are substantially different up to a number of eight. At this point, two similar patterns appear. This suggests eight as an appropriate

number of clusters to retain, which is in good concordance with the eight PPs found by the manual classification of Soler (1977) and the seven of Guardans and Palomino (1995).

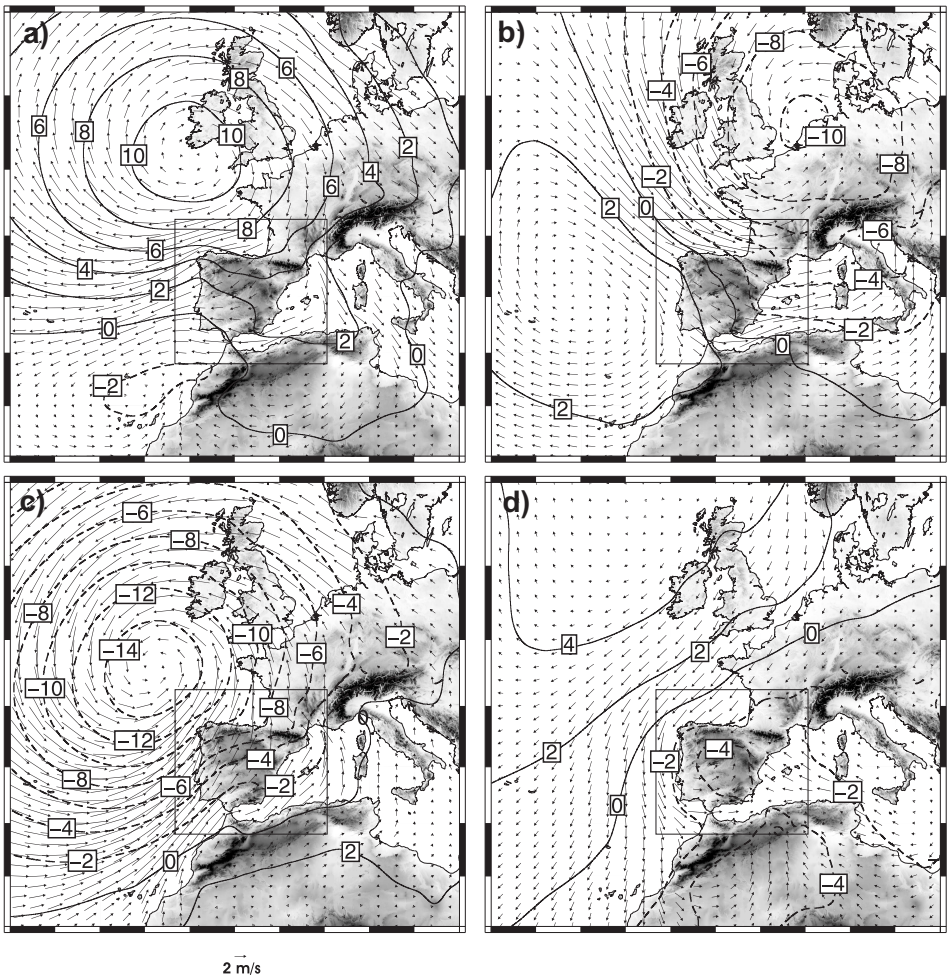

Fig. 6.16: Mean anomaly SLP fields (contour lines) and mean anomaly wind vectors at 10 m above ground level for days classified as: PP1 (a), PP2 (b), PP3 (c), PP4 (d), PP5 (e), PP6 (f), PP7 (g) and PP8 (h). The spatial window in which the classification is performed is also displayed (interior rectangle).

The mean anomaly fields of the resulting eight PPs are displayed in Figure 6.16. The classification was performed for a geographical window covering mainly the IP. However, the PPs are shown for a larger window for a better representation of the synoptic

features. The eight PPs are essentially different and are consistent with those PPs found by other authors in previous SLP classifications over the IP (Soler, 1977; Guardans and Palomino, 1995; Zhang et al., 1997; Trigo and DaCamara, 2000). The monthly frequencies of occurrence of the PPs are displayed in Figure 6.17. Some of the patterns show a clear annual wave. For instance, PP4 and PP8 present a maximum during summer and PP3 and PP5 during winter. For a clearer physical interpretation of the PPs, the mean SLP and the mean wind vectors at 10 m above ground level for the days belonging to each PP are displayed in Figure 6.18.

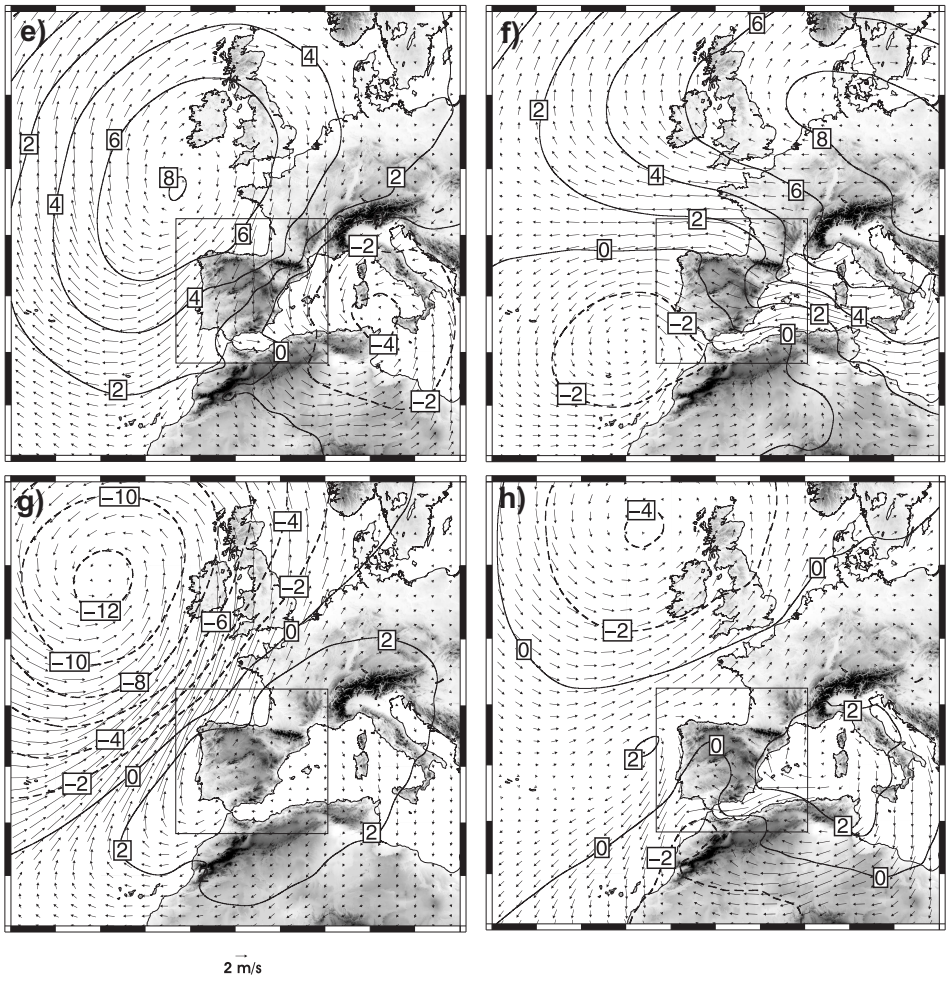

Fig. 6.16: Continuation.

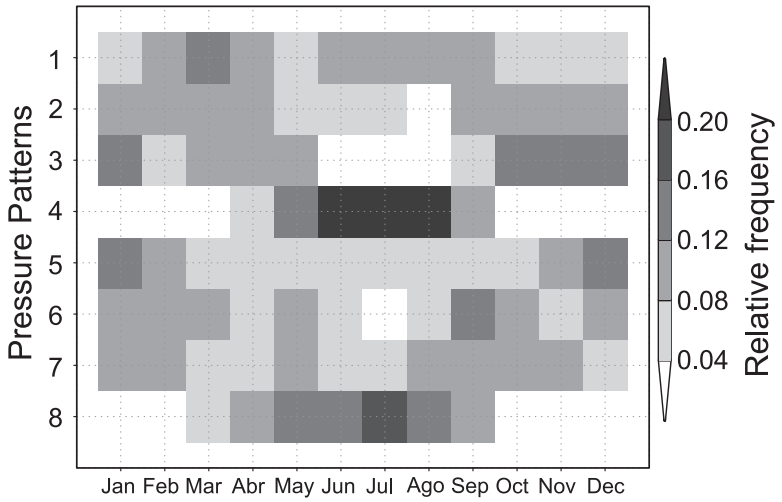

Fig. 6.17: Monthly frequency of the eight PPs (Fig. 6.16).

The PP1 presents positive SLP anomalies in the North of the IP and negative in the Northwest of Africa (Fig. 6.16a) as a consequence of the stationary Azores high pressure center displacement towards western Europe (Fig. 6.18a). PP2 shows negative SLP anomalies in central Europe and slightly positive over the Azores High (Fig. 6.16b) allowing for northwestern flows over the IP (Fig. 6.18b). A strong negative anomaly center in the Northwest of the IP dominates the PP3 (Fig. 6.16c). It represents the low pressure systems which intercept the IP coming from the polar front (Fig. 6.18c). The low pressure systems are more frequent at the late Autumn and during Winter, with a secondary maximum in Spring (Fig. 6.17). This can be considered a consequence of the lower latitudes that the Azores high visits in Winter, which allows that perturbations from the polar front reach the IP more frequently (Zimmerschied, 1949). The PP4 presents negative anomalies centered at the IP and in the North of Africa (Fig. 6.16d). It corresponds to the typical summer pattern (Fig. 6.17) in which the Azores high pressure center reaches higher latitudes and extends towards central Europe (Fig. 6.18d). These negative anomalies are related to the thermal lows over the IP and northern Africa originated by the strong ground heating (Font, 2000). This pattern appears in manual classifications (Soler, 1977; Guardans and Palomino, 1995) though it has been reported that certain automated SLP classifications, different to the eingenvector methodology herein used, have usually problems in its detection (Martín-Vide, 2001). The PP5 shows negative anomalies over the Mediterranean sea and positive in western Europe (Fig. 6.16e). The PP6 presents positive anomalies over eastern Europe and negative over the position of the Azores high (Fig. 6.16f). It is associated with the backward movement of the Azores high pressure center in benefit of the winter Siberian high (Fig. 6.18f). The PP7 reveals positive pressure anomalies over the IP which are also extended towards central Europe and block the perturbations coming from the polar front (Fig. 6.16g). It represents a NE extension

of the Azores high pressure center (Fig. 6.18g). Finally, the PP8 shows weak anomalies over the IP with a positive center over the Mediterranean sea (Fig. 6.16h). This leads to a typical situation where the Azores high blocks western circulations and with high pressures over the Mediterranean sea (Fig. 6.18h).

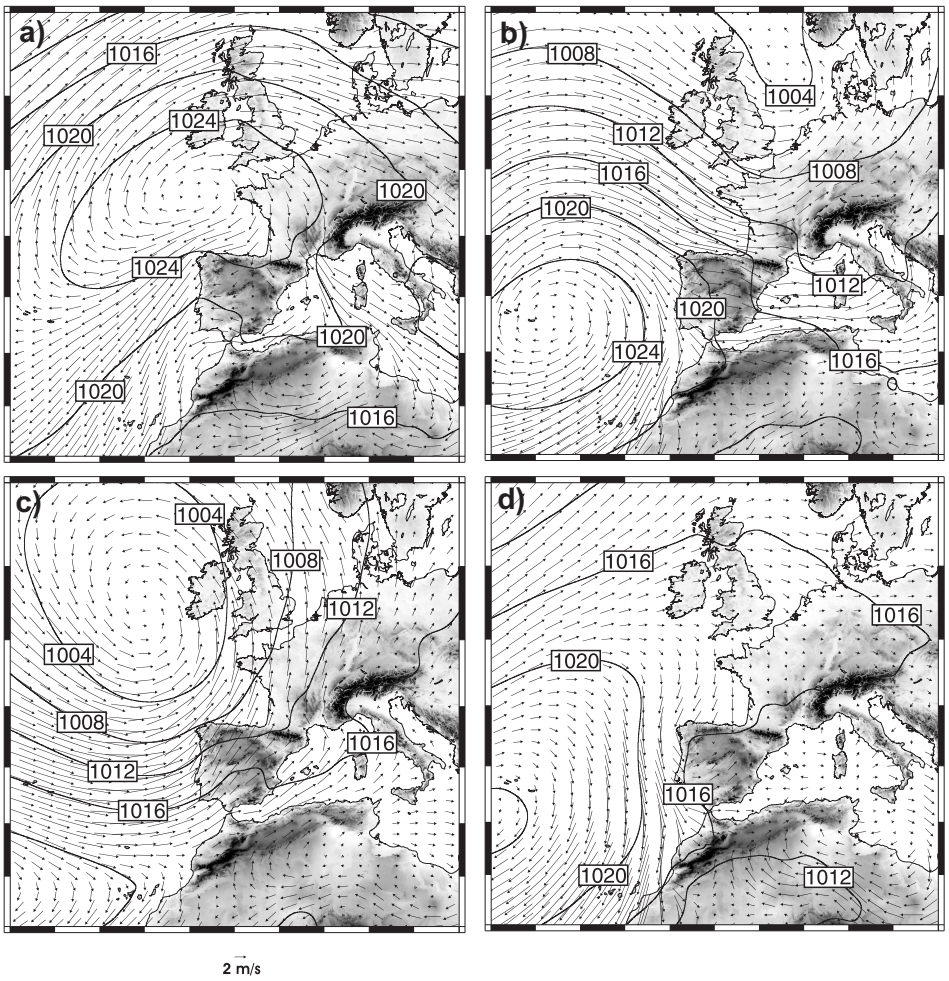

Fig. 6.18: Mean SLP (contour lines) and mean wind vectors at 10 m above ground level for days classified as a) PP1, b) PP2, c) PP3, d) PP4, e) PP5, f) PP6, g) PP7 and h) PP8.

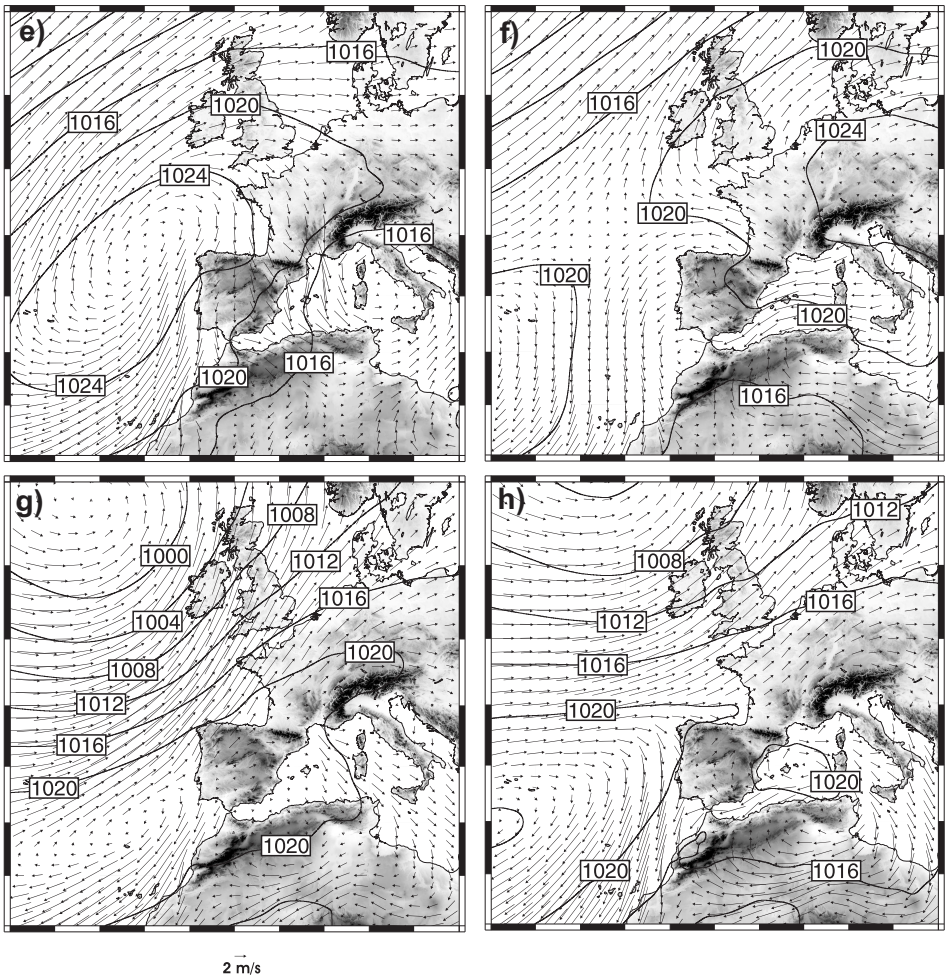

Fig. 6.18: Continuation.

6.6 Relationship between the synoptic scale and surface circulations

The relative frequency of the PPs (Fig. 6.16) is summarized in Table 6.2 (first column). The associations between the WPs (Fig. 6.7) and the synoptic patterns are also displayed in Table 6.2 (column 2-7). They are calculated by dividing the number of days that a certain WP appears in each particular PP by the total number of days within that PP. Therefore, it represents the conditional probability of the WPs given the occurrence of a particular PP. Table 6.2 shows that several WPs present a high association with a

specific PP. For instance, WP3 and WP1 are strongly associated with PP2 (62 %) and PP5 (76 %), respectively. Additionally, it is posible to find a situation in which various WPs are generated by a particular PP (e. g. PP7) or alternatively a certain WP can be produced under several PPs (e. g. WP1).

Table 6.2: PPs (Fig. 6.16) relative frequency and their associations with each WP (Fig. 6.7). Associations superior to 0.20 appear in bold.

	Frequency (%)	WP1	WP2	WP3	WP4	WP5	WP6
PP1	16.7	**0.50**	**0.38**	0.03	0.08	0.01	0.00
PP2	16.5	0.18	0.02	**0.62**	0.02	0.02	0.14
PP3	15.2	0.01	0.01	0.05	0.15	**0.52**	**0.26**
PP4	12.9	**0.27**	**0.48**	0.01	0.15	0.06	0.01
PP5	10.8	**0.76**	0.09	0.14	0.01	0.00	0.00
PP6	10.0	0.03	0.15	0.02	**0.43**	**0.36**	0.01
PP7	9.6	0.13	0.16	**0.21**	**0.22**	0.11	0.17
PP8	8.2	0.13	**0.35**	0.04	**0.40**	0.07	0.01

The next paragraphs will briefly describe the main associations between the PPs and the WPs (bold in Table 6.2). The discussion will be based on the different physical properties, temperature and moisture, of the air masses that affect the IP at different times of the year.

6.6.1 High pressures to the north of the Iberian Peninsula

PP1 contributes to anticyclonic geostrophic circulations over the CFN that due to the ageostrophic balance display a northern component (Fig. 6.18a). The resulting circulations are channeled by the mountain systems that surround the CFN producing northwestern circulations on the Ebro valley (Fig. 6.18a). This structure reveals the highest association with the northwestern WP1 (50 %, see Table 6.2) configuration (Fig. 6.8). The northwestern wind anomalies (Fig. 6.8a) show slightly negative wind speed anomalies in the North part of the CFN changing to positive towards the South and, increasing their values down the Ebro valley (Fig. 6.8b); the anomalies of the re-analysis data also increase their values along the Ebro valley but in lower magnitude than the observational values. The observed wind speed intensification is due to the higher positive anomalies in the headboard of the Ebro valley compared to those in the valley mouth close the Mediterranean sea (Fig. 6.16a). This configuration enhances the pressure gradient along the valley and intensifies the winds (García, 1985). The northern circulations (Fig. 6.18a) introduce cold air into the IP originating negative temperature anomalies over the whole CFN and of higher magnitude in the North than in the South adequately reproduced by the analysis (Fig. 6.8c). The advected air produces positive moisture anomalies as a consequence of its maritime origin that seem to be blocked by the *northern*, *western* and *eastern mountains* leading to negative anomalies in the South of the CFN (Fig. 6.8d); the analysis reproduce this structure although, as expected, they underestimate the magnitude of the observations. The global radiation anomalies show negative (positive) values

in the North (South) indicative of cloudiness (clearer skies) in concordance with the analysis (Fig. 6.8e). In spite of the covered skies and positive moisture anomalies in the North of the CFN (Fig. 6.8d,e), the precipitation anomalies show general negative values (Fig. 6.8f), which suggests that the expected precipitation would be of low intensity.

PP1 also shows important associations with WP2 (38 %). This WP also presents northwestern wind anomalies over the CFN (Fig. 6.9a) although of weaker intensity than WP1 (Fig. 6.8a) as shown also in Figure 6.6a,b. The prevailing association with WP1 or WP2 depends largely on the intensity and possition of the high pressure center (see Fig 9.1 in Appendix A). A more eastward displacement or zonal orientation produces a blockage of the northern circulations and thus weaker wind speeds over the CFN (Fig. 6.18a), and higher associations with WP2 should be expected; whereas a westward displacement or meridional orientation favors the northern circulations to the CFN along the eastern side of the high pressure center (Fig. 6.18a) which should lead to higher associations with WP1. In addition, a strengthening (weakening) of the high pressure center would increase the pressure gradient along the Ebro valley and higher associations with WP1 (WP2) would be expected. PP1 appears during the whole year with its maximum of occurrence during March and important frequencies during Summer (Fig. 6.17). This is in concordance with the apparence of WP1 during the whole year and the maximum frequency that WP2 displays during the summer (Fig. 6.14). A more detailed explanation of the WP2 properties will be further introduced when discussing the PP4, which shows the highest association with WP2 (Fig. 6.14).

6.6.2 Northwestern flows

PP2 displays a strong association with WP3 (62 %). This WP presents northwestern wind anomalies (Fig. 6.10a) with negative wind speed anomalies over the CFN especially in the region between the *western* and *eastern mountains*, although some positive anomalies appear at mountain stations (Fig. 6.10b). These stations are exposed to northwestern circulations which are dominant in this WP (Fig. 6.10a). The wind speed anomalies of the ECMWF data for this pattern reveal high positive values over the sea near the CFN coast and a sudden decrease down the Ebro valley (Fig. 6.10b). This could be interpreted as a consequence of the blocking effects originated by the Cantabrian Mountains and the Pyrenees presence (Fig. 1.2). Furthermore, the isobars run along the axis of the Ebro valley (Fig. 6.18b) which produces a weak local pressure gradient and so winds are not intensified along the valley. WP3 involves large positive precipitation anomalies in the northern part of the CFN that decrease towards the South (Fig. 6.10f). This structure is reproduced by the ECMWF data, although the precipitation gradient is weaker and penetrates inside the Ebro valley. As a consequence, the relative humidity content in the atmosphere over the North of the CFN is larger than in the South (Fig. 6.10c). This gradient is also observed in the analysis data, although it underestimates the magnitude. In addition, the general negative global radiation anomalies in concordance with the analysis (Fig. 6.10e) are an indication of cloudiness. The temperature anomalies are negative in both the observations and the analysis (Fig. 6.10c). In the view of PP2 (Fig. 6.18b), this atmospheric state could be considered as the effect produced by an originally polar maritime air mass that in its displacement along the Atlantic ocean

increases its water content, producing precipitation when it rises up the *northern* and *western mountains* of the CFN (García and Reija, 1994). Therefore, the properties of this WP3 (Fig. 6.10) are similar to the ones of WP1 (Fig. 6.8) but the more zonal orientation of the flow introduces less colder air which allows for larger water content; thus, it produces higher precipitation amounts in the North of the CFN and cloudiness and moisture penetrate deeper inside the Ebro valley. PP2 appears during the whole year with less frequency during Summer (Fig. 6.17) in concordance with the WP3 occurrence.

6.6.3 Northwestern low pressures

PP3 typically generates *Bochorno* situations characterized by warm moist winds blowing up the Ebro valley (García, 1985). It presents high associations with WP5 (52 %). This WP presents southeastern anomaly winds (Fig. 6.12a) due to the cyclonic circulations and the ageostrophic balance towards the low pressure center. There is also an important pressure gradient along the Ebro valley which intensifies the southeastern circulations and thus, the wind speed anomalies are positive (Fig. 6.12b). The ECMWF wind speed anomalies show an increasing gradient up the Ebro valley with slight positive values over the CFN (Fig. 6.12b). The southeastern flow introduces warm moist air from the Mediterranean sea which produces positive anomalies of temperature and relative humidity in the Ebro valley (Figs. 6.12c,d). The temperatures are higher and the air is drier at the North of the CFN due to the Foehn effect associated with the *eastern* and *western mountains*. This is reproduced by the analysis data, although it is somewhat displaced with respect to observations. The global radiation anomalies are negative (Fig. 6.12e) and the precipitation anomalies positive (Fig. 6.12f) except in the norhtern valleys of the Pyrennes. It indicates the presence of cloudiness with precipitation in the South of the CFN and clearer skies in the North in concordance with the Foehn effect mentioned above. The highest positive precipitation anomalies appear at the *western mountains* probably due to the mild moist air forcefully lifted at the end of the Ebro valley. Both, the global radiation and precipitation gradients along the CFN are reproduced by the analysis data (Fig. 6.12e,f). The maximum frequency that this PP shows during the winter (Fig. 6.17) involves a maximum frequency of occurrence of WP5 during this season (Fig. 6.14).

The PP3 is also the main responsible of WP6 (26 %). This WP shows southwestern wind anomalies over the CFN (Fig. 6.13a) in concordance with the cyclonic circulations displayed over the IP by PP3 (Fig. 6.18c). The flows present negative wind speed anomalies that are stronger in the Southeast part of the CFN with a similar structure observed in the analysis data although without reach negative values (Fig. 6.13a). The negative anomalies seem to be a consequence of the blocking effect produced by the Iberian system to the general flow. The temperature anomalies present positive values over the CFN posibly as a consequence of the Foehn effect produced by the Iberian system (Fig. 6.13c). This WP shows a less organized relative humidity structure with general negative anomalies in the region but, with sporadic locations with positive values (Fig. 6.13d). The global radiation anomalies are weak with positive values in the Norhtwest of the CFN and negative in the Southwest (Fig. 6.13e). The precipitation anomaly filed shows negative values with a few exceptions in the *northern mountains* (Fig. 6.13f). This WP6 also

displays its maximum of occurrence in winter (Fig. 6.14) in concordance of the maximum frequency of occurrence of the PP3 (Fig. 6.17).

The associations of PP3 with WP5 are however more frequent (52 %) than with WP6 (26 %). The apparence of one or the other could be explained in terms of the relative intensity and position of the low pressure system that appears in PP3 (Fig. 6.18c and Fig. 9.3 in Appendix A). If the location of the low pressure system leads to a more southward geostrophic flow over the IP, PP3 would produce higher associations with the southeasstern circulations of WP5 due to the ageostropic balance which introduces the flow towards the low pressures and thus favours the southeastern circulations along the Ebro valley (Fig. 6.12a). The circulations are intensified by the strong pressure gradient along the valley, leading to the high winds shown by WP5 (Fig. 6.12b). If the low pressure system location leads to southwestern flows over the IP with enough intensity to pass over the Iberian system, PP3 could produce southwestern circulations over the CFN as shown in WP6 (Fig. 6.13a). However, if the southwestern circulations do not present enough intensity to pass over the Iberian system, this situation could evolve in a similar way as the one described before leading to southeastern circulations over the CFN and thus associations with WP5. This interpretations is also in concordance with the higher frequency of occurrence of the southeastern circulations of WP5 over the southwestern ones of WP6 under this PP3.

6.6.4 The thermal low

The cyclonic circulations developed around the IP thermal low favors the flow towards the interior of the peninsula due to the ageostrophic balance (Font, 2000). This tendency can be appreciated in the ECMWF data (Fig. 6.18d). The resulting inland circulations are channeled by the local CFN orography and therefore this PP presents high associations with the northwestern WP1 (27 %) and WP2 (48 %). The appearance of WP1 instead of WP2 would depend on the specific position and intensity of the Azores high as was commented in relation to PP1 (see Fig 9.4 Appendix A). The highest associations are with WP2 that shows weak northwestern wind anomalies over the CFN (Fig. 6.9a) and general negative wind speed anomalies (Fig. 6.9b). The analysis data reveals that the low wind speed anomalies are a general characteristic of the flows over the whole IP (Fig. 6.9b). The northern component of the flow together with its maritime origin introduces slightly colder and moister air than usual (Figs. 6.9c,d) which is blocked by the *western* and *northern mountains*. As a consequence, the air is warmer and drier than usual over the central and especially southern regions of the CFN. The temperature (moisture) anomalies of the analysis data show a maximum (minimum) centered over the IP and a decrease towards the coasts where near zero values are reached. This in concordance with the temperature (moisture) gradient along the CFN. The global radiation presents positive anomalies over the CFN (Fig. 6.9e) while the ECMWF data again reveals that this is a characteristic over the whole IP, suggesting cloudless skies in concordance with the negative anomalies of precipitation (Fig. 6.9f). Both characteristics are associated with a subsidence inversion, which together with the low moisture air prevent the formation of clouds even if the radiation anomalies reach high values (Zimmerschied, 1949). This

PP4 shows a clear maximum centered on the summer (Fig. 6.17) as its associated WP2 (Fig. 6.14).

6.6.5 Northwest-Southeast pressure gradient

This situation produces a certain meridional orientation of the Azores high pressure center that introduces north-northwestern flows over the CFN due to the ageostrophic balance (Fig. 6.18e). This is the typical PP producing *Cierzo* winds, characterized by its cold and dry strong northwestern circulations in the Ebro valley (García, 1985). It shows a clear relationship with the WP1 (76 %). This WP1 and its associated atmospheric state (Fig. 6.8) were already discussed in relation with PP1. The wind speed intensification on the Ebro valley (Fig. 6.8b) is a consequence of the strong pressure gradient along the valley produced by the positive and negative SLP anomaly centers (Fig. 6.18e) whereas the negative temperature anomalies are produced by the meridional orientation of the Azores high pressure center which leads to advection of cold air from the North (Fig. 6.8c). The humidity and cloudiness are blocked by the mountains in the North of the CFN leading to the dry winds observed in the Ebro valley (Fig. 6.8d). PP5 appears during the whole year (Fig. 6.17) in concordance with the WP1 ocurrence (Fig. 6.14).

6.6.6 Siberian high pressures

The PP6 appears with its maximum frequency in Winter and Autumn although it is somewhat present during the whole year (Fig. 6.17). This PP can advect cold air from the Siberian polar continental air mass but the Pyrenees acts as a blocking barrier and thus such northern flows are very infrequent (Table 6.2). The highest associations are found with the southeastern circulations in WP4 (43 %) and in WP5 (36 %) which advect warm (Figs. 6.11c and 6.12c) and moist (Figs. 6.11d and 6.12d) air from the Mediterranean sea. The main difference between WP4 and WP5, apart from the wind speed (Figs. 6.11b and 6.12b), is the associated precipitation and global radiation. The WP4 precipitation anomalies are slightly positive in the Ebro valley and negative in the rest of the CFN (Fig. 6.11f) while WP5 shows a similar structure but with stronger positive precipitation anomalies that also affects central areas of the CFN (Fig. 6.12f). Acordingly, the WP4 global radiation anomalies are slightly negative over the Ebro valley changing to positive in the *western* and *eastern mountains* (Fig. 6.11e) and, the WP5 anomalies are negative excepts to the North of the *northern mountains* wherein they are slightly positive (Fig. 6.12e). This could be related to the WP5 higher than WP4 wind speeds (Fig. 6.11b and 6.12b) which transport cloudiness with precipitation from the Mediterranean sea deeper inside the Ebro valley.

6.6.7 Northeastern extension of the Azores high

On one hand, this pattern can introduce subtropical maritime air masses into the IP. On the other, important low pressure anomalies are located in the Northwest of the IP which could potentially affect the northern regions and thus, the CFN. It does not show

a clear relationship with any WP (Table 6.2). This suggests that surface circulations are sensitive to variations of the above described combined structure, or even that upper level circulations play a more important role than the weak surface pressure gradient typical of these anticyclonic situations. The highest associations are found with WP3 (21 %) and WP4 (22 %). PP7 can appear during the whole year although it is more frequent in Autumn and Winter (Fig. 6.17). The higher frequency of this PP during winter could be related with the known tendency of high pressures to develop over cold areas due to its strong stability, as is the case of continents during the winter season. However, this PP also appears during summer, and a reason for this development seems to be related with the presence of warm air at upper levels over the IP which produces strong stability (Font, 2000).

6.6.8 Anticyclonic situation

The PP8 produces weak pressure gradients over the IP and in particular along the Ebro valley which results in moderate wind speeds over the CFN. Thus, the most important associations are found with WP2 (35 %) and WP4 (40 %), with the weakest wind speeds (Table 6.1). These WPs present opposite surface flows. It could be argued that the circulation sense of both WPs (NW for WP2 and SE for WP4) is influenced by the relative intensity of the high pressure center over the Mediterranean region with respect to the high pressures at the headboard of the Ebro valley (see Fig. 9.8 Appendix A). This PP is more frequent during Summer (Fig. 6.17) like WP2 and WP4 which also dislplay their maxima of ocurrence on this season (Fig. 6.14).

6.7 Conclusions

Two methodologies, one based on the spatial similarity of wind fields and the other on its temporal variance were used to classify the mean daily wind fields over the CFN. Both methodologies produce similar WPs. The classification based on the spatial similarity was selected and a total of six WPs were identified (Fig. 6.7). It was found that the northwestern winds are the dominant pattern, followed by the southeastern circulations. The NW-SE is the main direction of the mountain ridges that surround the CFN and thus, it is evidenced the strong influence of the orography effects over the surface circulations.

To understand the influence of CFN surface circulations on its regional climate, temperature, relative humidity, global radiation and precipitation patterns corresponding to each WP were also analyzed. The WPs present different advective regimes; in particular, the cold dry *Cierzo* and the warm moist *Bochorno* were recognized. A surface WP responsible for the high precipitation in the north part of the CFN was also recognized.

A SLP map-pattern classification was also performed. Eight PPs were identified and related to the surface WPs. The relations found between the PPs and the WPs were consistent enough to understand the basic forcing mechanisms of the different regional circulations and their influence on the CFN regional climate. The pressure gradients along the Ebro Valley and the ageostrophic balance seem to successfully describe the general characteristics of the surface flows.

The strong influence of the stationary subtropical high pressure center over the Azores islands in the CFN circulations has been shown. Its location at higher (lower) latitudes blocks (allows) the penetration of perturbations from the polar front, controlling the CFN surface flows and their physical properties.

7

Climatology: WRF evaluation*

This makes it quite clear how far removed we are from a complete understanding of the motion of fluids and that my exposition is no more than a mere beginning. Nevertheless, everything that the Theory of Fluids contains is embodied in the two equations formulated above (S 34), so that it is not the laws of Mechanics that we lack in order to pursue this research but only the Analysis, which has not yet been sufficiently developed for this purpose. It is therefore clearly apparent what discoveries we still need to make in this branch of Science before we can arrive at a more perfect Theory of the motion of fluids.

L. Euler, 1755

Principes généraux du mouvement des fluides. Académie Royale des Sciences et des Belles-Lettres de Berlin, Mémories, 11, 274-315.

The realism of the WRF dynamical downscaling in reproducing areas of coherent variability of surface wind over the CFN has been shown in Chapter 5. It is also interesting to evaluate the capability of the downscaling approach to reproduce the climatological surface flows over the region. This is the topic of the present chapter.

The performance of the downscaling to reproduce the long term climatological flow is analyzed in a first step. The discussion highlights the effects of the orography prescribed in the WRF model over the downscaled wind field. Then, the performance of the downscaling to reproduce the climatological flow of the six WPs identified in the Chapter 6 is analyzed. In order to provide a detailed evaluation of the downscaling capability to reproduce the influence of the orography, the analysis over each WP is broken down into the four wind subregions with different wind variability identified with observations in Chapter 4 (Fig. 4.11a). Special emphasis will be paid to evaluate the performance of the downscaling in reproducing the surface wind field as a function of the large scale synoptic situation. With this aim, the downscaling capability to reproduce each WP will be evaluated separately for each one of its associated PPs (Table 6.2).

* The main contents of this chapter are included in:
paper prepared for sumbision (reference to be updated in the last version of this manuscript)

7.1 Background and motivation

A dynamical downscaling demands high computational resources. One always need to carefully design the experiment in order to dimension the available resources. The spatial resolution achieved and the length of the simulated period are usually important factors to take into consideration. Certain studies perform a high-spatial resolution simulation during the few days of an intensive observational campaign (e.g. Zhong and Fast, 2003). The availability of measurements allows for a detailed evaluation of the simulation performance to reproduce the specific regional-scale process under study (e.g. Georgelin and Richard, 1996; Bonnardot and Cautenet, 2009) or to test sensitivities to the model configuration (e.g. Zhang and Zheng, 2004; Michelson and Bao, 2008). For instance, Masson and Bougeault (1996) used data from an intensive observation period of the Pyrénées Experiment (PYREX) to study the dynamics of a *Cierzo* situation. Other case studies, analyze the influence of different synoptic situations on the regional-scale circulations (Soriano et al., 2006; Mengelkamp, 1999; Källstrand et al., 2000; Yang et al., 2008).

In certain situations is desirable to use a coarser spatial resolution in order to simulate a longer temporal period. Such is the case of regional climate simulations that evaluate the wind resource over a given area for the present Walter et al. (2006); Jiang et al. (2008) or for future climate change scenarios (Debernard et al., 2002; Pryor et al., 2005c; Debernard and Røed, 2008).

An alternative approach that allows to obtain high-resolution information of the climate over a given region are the statistical-dynamical methodologies (Mahrer et al., 1985; Fuentes and Heimann, 1996; Frank and Landberg, 1997). In such approaches, a climatological classification is performed, and the mesoscale model is run for representative situations of each resulting group; the final climatology is the sum of the individual simulations weighted by the frequency of apparition of its group (Frey-Buness et al., 1995; Fuentes and Heimann, 2000). Since the mesoscale model is run for a few situations of each group, it can be run at high horizontal resolution. Further reduction of the computational resources can be obtained by performing the mesoscale simulations in a dynamic mode (Pielke, 1985) in which the thermal effects are not considered (Wippermann and Gross, 1981; Heimann, 1986; Brücher et al., 1994; Sandström, 1997; Mengelkamp et al., 1997; Mengelkamp, 1999). However, this last approach has shown to be somewhat problematic (Egger, 1995) and the simulations including the thermal effects to deliver a superior performance (Heimann, 2001).

The present research evaluates the ability of the WRF simulation (Chapter 3) to reproduce the climatological WPs identified over the CFN (Chapter 6). Hence, thermal effects are condidered in the simulation and the statistical-dynamical downscaling is avoided. The high resolution (2 km) and the reasonably long simulated period (over thirteen years) allows for a precise evaluation of the model performance. The evaluation focuses on the influence that the large scale and the smoothed orography used in the simulation exerts on the simulated wind field. To our knowledge there are no previous studies that analyzed a simulation with the characteristics, in terms of spatial resolution and length of the simulated period, of the one used in the present assessment.

7.2 Data preparation

The observed wind at the 41 sites that define the WPs is represented by the simulations at the grid points co-located to the observational sites. This representation is the same one that was adopted in Chapter 5 to analyze the simulation performance in reproducing the wind variability. Some limitations of this representation were discussed in that chapter. The daily mean wind simulated at 10 m above ground level is again used in the comparison, since the majority of the wind sensors are situated at this height (see Table 2.1). The focus now is to analyze the climatological flow and thus, the mean of the time series is important. Hence, the few wind records taken at 2 m above ground level (seven stations) were extrapolated to 10 m. The extrapolations were performed using the power law with an exponent of 1/7 as it was done for instance in Pryor et al. (2005a). As a final step in the data preparation, the simulated counterparts of missing daily observations were removed from the 41 simulated time series in order to perform the evaluation with equivalent datasets.

7.3 Wind field reproducibility: on the effects of topographical features

This section evaluates the capability of the WRF dynamical downscaling to reproduce the climatological wind field. The assessment is accomplished in two steps. Some initial insight into the simulation performance is obtained by analyzing the long term climatology (Section 7.3.1). The analysis stresses the influence that the smoother orography used in the WRF simulation exerts on the simulated wind. A more detailed evaluation is accomplished by analyzing the reproducibility of the typical surface flows over the area, the WPs (Section 7.3.2). This evaluation shows the effects that a large scale misrepresentation produces on the surface wind field. Section 7.4 provides a more in depth evaluation of the influence that the large scale exerts on the downscaling ability to reproduce the climatological surface circulations.

7.3.1 Reproducibility of the long term climatological flow from 1992 to 2005

The wind speed bias and the root mean squared error (RMSE) of the wind direction calculated with each pair of simulated and observed time series are displayed in Figure 7.1. The wind direction is better reproduced at the mountain tops (stations 3, 4, 16, 20, 35 and 37, Fig. 1.1) than at the valleys likely due to a weaker influence of the surrounding topography on the flow at the former locations. There is a tendency of the downscaling to underestimate the wind speed at the observational sites located at mountain tops, which are the windiest ones, and to overestimate it at the locations situated in the valleys, the less windy areas. This causes a reduced spatial variability of the downscaled wind speed field compared to the observational one. To illustrate this result, the mean wind speed is represented against its standard deviation for both observations and simulations at each station (Fig. 7.2). The mean and the standard deviation are correlated, thus locations with a high mean wind speed also present a high standard deviation. This is

typical of positively defined variables (see Section 4.3). The simulated wind reproduces this relationship between mean and standard deviation, but the range of variability is considerably smaller than the observed one.

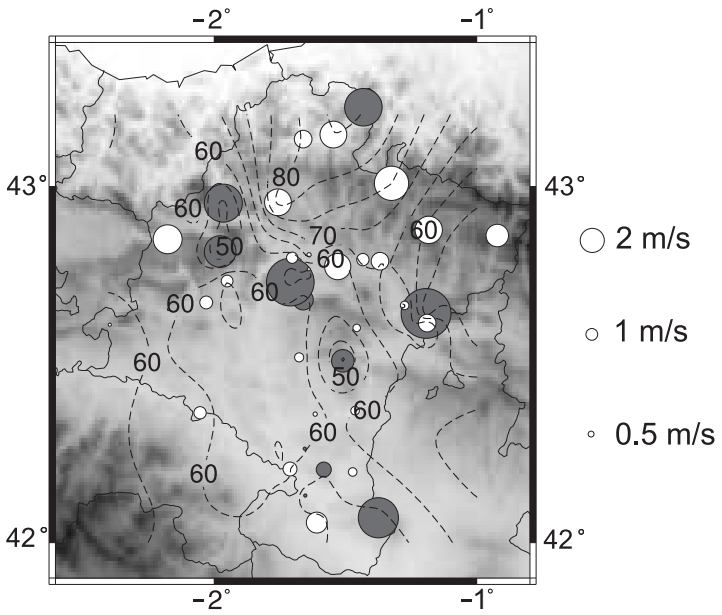

Fig. 7.1: Wind speed bias (circles) and RMSE of the direction (contour lines) calculated with the observed and simulated time series. The wind speed bias is defined as the mean observed wind speed minus the mean simulated wind speed. The radius of the circles is proportional to the magnitude of the bias and the white (gray) color denotes negative (positive) values.

The different behavior that the simulation shows in reproducing the climatological wind speed at valleys and mountain locations could stem from the smoother topography within the mesoscale model. It could be argued that even the high spatial resolution used, 2 km, is not able to reproduce with enough accuracy the complexity of the terrain over the CFN; not to mention the influence that the terrain variability within a simulated grid cell could have in the simulation. The misrepresentation of the orography can potentially introduce systematic errors in the simulation which ultimately would lead to the wind speed underestimation (overestimation) at mountain tops (valleys). In order to explore this possibility, a more realistic terrain dataset of 90 m of resolution (Farr et al., 2007) is compared to the topography used in the WRF simulation. The differences between these two datasets are displayed in Figure 7.3a. The horizontal resolution of 2 km used in the WRF model seems to be adequate over the relatively flat terrain of the Ebro valley, but it is less accurate in its representation of the higher complexity of the terrain in northern areas of the CFN.

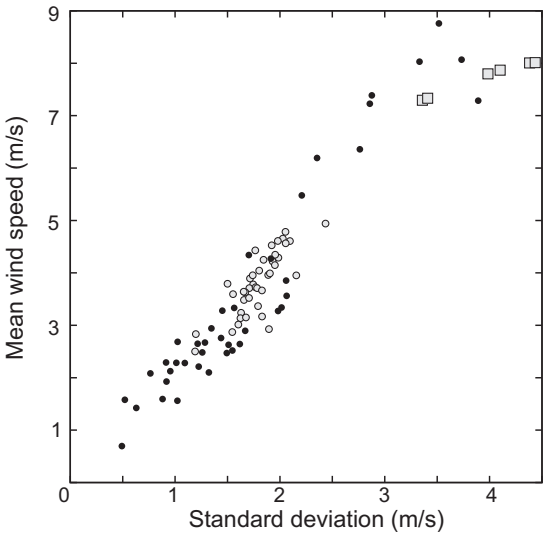

Fig. 7.2: Scatter diagram of the mean against the standard deviation of the 41 observed (dark circles) and simulated (light circles) wind speed time series. The scatter diagram calculated with the simulated wind at the grid points co-located near the actual height of the mountain top stations are also represented (squares).

On the basis of the previous arguments, a plausible explanation for the wind speed underestimation at the mountain tops could be attributed to the representation of flatter mountains than actual ones by the mesoscale model. Even at the high horizontal resolution used (2 km), the mountains are considerably smoother than reality (-452, -521, -277, -371, -137 and -141 m for the stations 3, 4, 16, 20, 35 and 37, respectively), and keeping in mind that the wind speed increases with height in the lower troposphere, this underestimation could be causing the wind speed underestimation at the top of the peaks. This can be better understood by representing the wind speed bias *versus* the topography differences of the terrain height used in the simulation (2 km) and the more realistic one presented above (90 m, Farr et al., 2007). For this purpose, the 90 m data is averaged within each grid cell. The comparison is displayed in Figure 7.3b, where it is also possible to appreciate the different behavior of the simulation at the four subregions over the CFN. The stations at mountain tops, sites within the MS subregion (squares in Fig. 7.3b), show a positive bias that increases as the terrain differences increase ($r = 0.77$). This suggests that the wind speed underestimation is accentuated with the underestimation of the mountains height. Therefore, the wind speed underestimation is at least partially associated with deficiencies in the terrain representation.

It also can be argued that mountain stations are well exposed to *quasi-geostrophic* winds, and therefore with a certain decoupling from surface effects. Hence, the wind speed at these locations could be represented by the simulated wind at the grid point co-

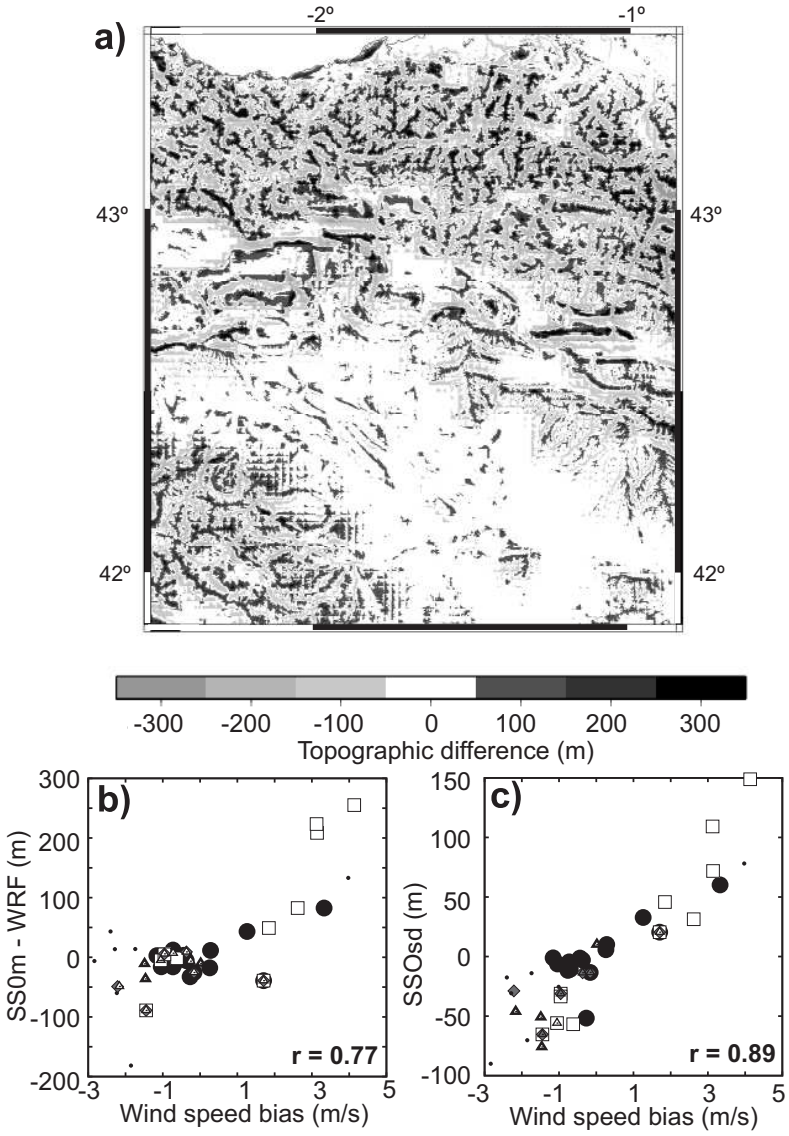

Fig. 7.3: (a) Topographic difference between the dataset of 90 m resolution (Farr et al., 2007) and the topography used in the WRF downscaling. (b) The difference between the average of the orography within each of the simulated volumes that encloses the locations of the observational sites and its associated value from the WRF simulation versus the wind speed bias. (c) The standard deviation of the subgrid-scale orography within each simulated volume that encloses the location of the observational sites versus the wind speed bias. The different symbols in panel b) and c) represent the different wind regions: EV (circles), MS (squares), NS (diamonds) and NV (triangles). The dots denote the stations that were not ascribed to any of the previous subregions.

located to the actual height of the station, instead of at 10 m above the simulated ground level. The influence of this change can be observed in Figure 7.2. The mean and standard deviation calculated with the simulated wind at the actual height of the stations are in a better agreement with the observed values than those obtained with the simulated wind at 10 m above ground. This result is an additional support to the hypothesis of an underestimation of the wind speed at mountain tops as consequence of the smoother topography used in the simulation.

A convincing explanation for the wind speed overestimation at the valleys is not straightforward. A plausible explanation could lie on the drag generated by the unresolved orography whose effects has been shown to be of relevance for the surface wind simulation (e.g. Beljaars, 2004; Rontu, 2006; Howard and Clark, 2007) and are not parametrized in the present simulation. According to this interpretation, the wind speed overestimation would be higher at the more complex terrain locations wherein the unresolved drag due to the subgrid-scale orography would be higher. This seems to be the case of the present simulation wherein the wind speed overestimation is higher at the sites located at the more complex terrain areas in the north of the CFN than at the smoother orography areas of the broad Ebro valley (see bias in Fig. 7.1). In order to analyze this hypothesis, the relationship between the wind speed bias and the standard deviation of the subgrid-scale orography is represented in Figure 7.3c. This standard deviation is calculated with the high resolution terrain dataset of 90 m (Farr et al., 2007) and provides an idea of the variability and thus, the complexity of terrain. Some relation between the standard deviation of the unresolved orography and the bias introduced by the WRF downscaling can be appreciated ($r = 0.89$). In particular, the sites located at valleys on the more complex terrain areas (triangles, NV; diamonds, NS; and dots) tend to show larger overestimations as the standard deviation increases. This points towards inaccuracies of the terrain representation as being at least partially responsible for the wind speed overestimation at the valleys.

7.3.2 Reproducibility of the wind patterns

This section explores the simulation performance to reproduce the typical surface circulations over the area (WPs). The mean observed/simulated wind field of days belonging to each one of the six WPs are displayed in Figure 7.4. In general, the downscaling reproduces the main characteristics of the climatological surface flows. The first three WPs display northwestern circulations (Fig. 7.4a,b,c). The spatial structure of WP1 and WP3 is better reproduced than the circulations of WP2. The following WP4 and WP5 display southeastern circulations with a better reproduction of the flow in the latter (Fig. 7.4d,e). The last WP6 shows southwestern circulations which are captured by the simulation (Fig. 7.4f). It is remarkable that the WPs that show an overall worse reproduction, WP2 and WP4, show a higher variability in the surface wind field than the other WPs.

A quantification of the skill displayed by the simulation to reproduce the climatological flow under each WP is break down for the four wind subregions with different wind variability (Fig. 4.11a) in Figure 7.5. The highest winds are observed at the subregion mainly formed by mountain stations (MS, Fig. 7.5a) which are systematically underestimated by the simulation (Fig. 7.5c). The wind speed at the valley subregions (EV, NS

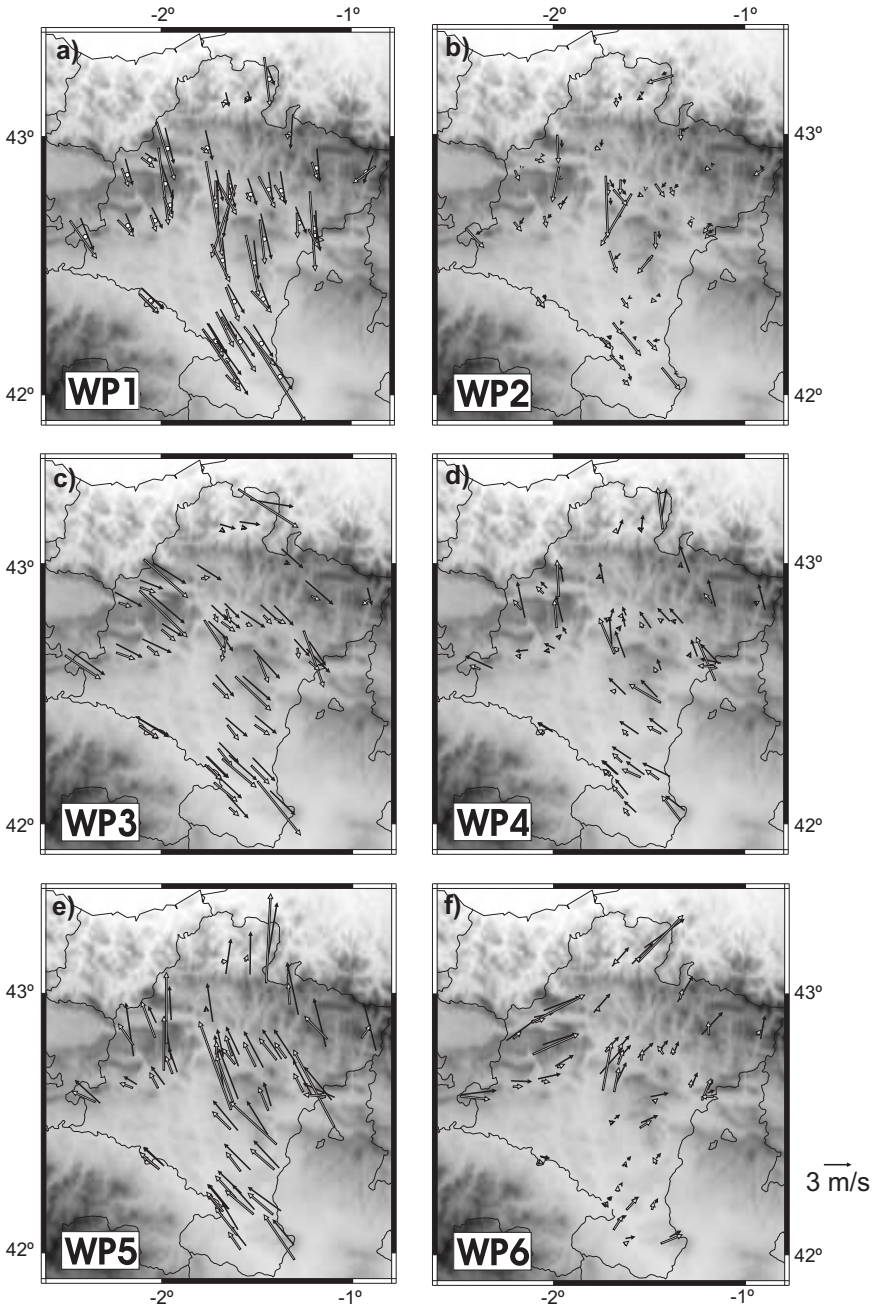

Fig. 7.4: Mean wind vectors calculated with the observations (white arrows) and simulations (black arrows) for days classified as a) WP1, b) WP2, c) WP3, d) WP4, e) WP5 and f) WP6.

and NV) is overestimated under WP3, WP4, WP5 and WP6, but there is a tendency to underestimate it also under WP1 and WP2 (Fig. 7.5c). At this point, it could be said that the long term tendency to overestimate the wind speed at the valleys noticed in the previous section (Fig. 7.1) is a consequence of the higher frequency of occurrence of WP3-6 (55.3 %) in comparison with that of WP1-2 (44.7 %). It will be shown in the next section that the wind speed underestimation at the valleys under WP1 and WP2 is a consequence of a misrepresentation at the large scale.

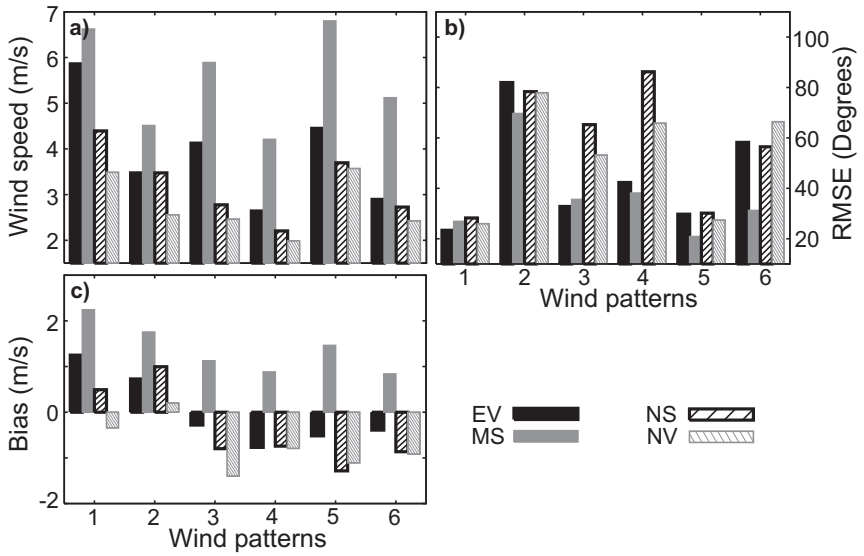

Fig. 7.5: Mean observed wind speed (a), RMSE of the wind direction (b) and wind speed bias (c) calculated with the observed and simulated regional time series for days belonging to each WP. The homogeneous wind subregions used to calculate the regional time series are defined in Figure 4.11a.

The MS subregion shows the best wind direction reproduction, with RMSE under 40 degrees except for WP2 under which the four subregions show poor scores (Fig. 7.5b). The circulations along the Ebro valley are also reasonably well captured as indicated by the moderate RMSE scores at the EV subregion under WP1, WP3, WP4 and WP5, the only exception being WP2 as indicated above. The simulation is less skillful in reproducing the cross Ebro valley winds of WP6. The other valley subregions, NS and NV, do not display a RMSE of the wind direction as low as in the EV, showing RMSE scores around 70 degrees, except for the strong circulations of WP1 and WP5 (Fig. 7.5a), under which the four subregions show RMSE below 30 degrees.

Additional understanding of the downscaling skill to reproduce the mean flow direction of the WPs can be obtained by comparing the observed and simulated wind roses

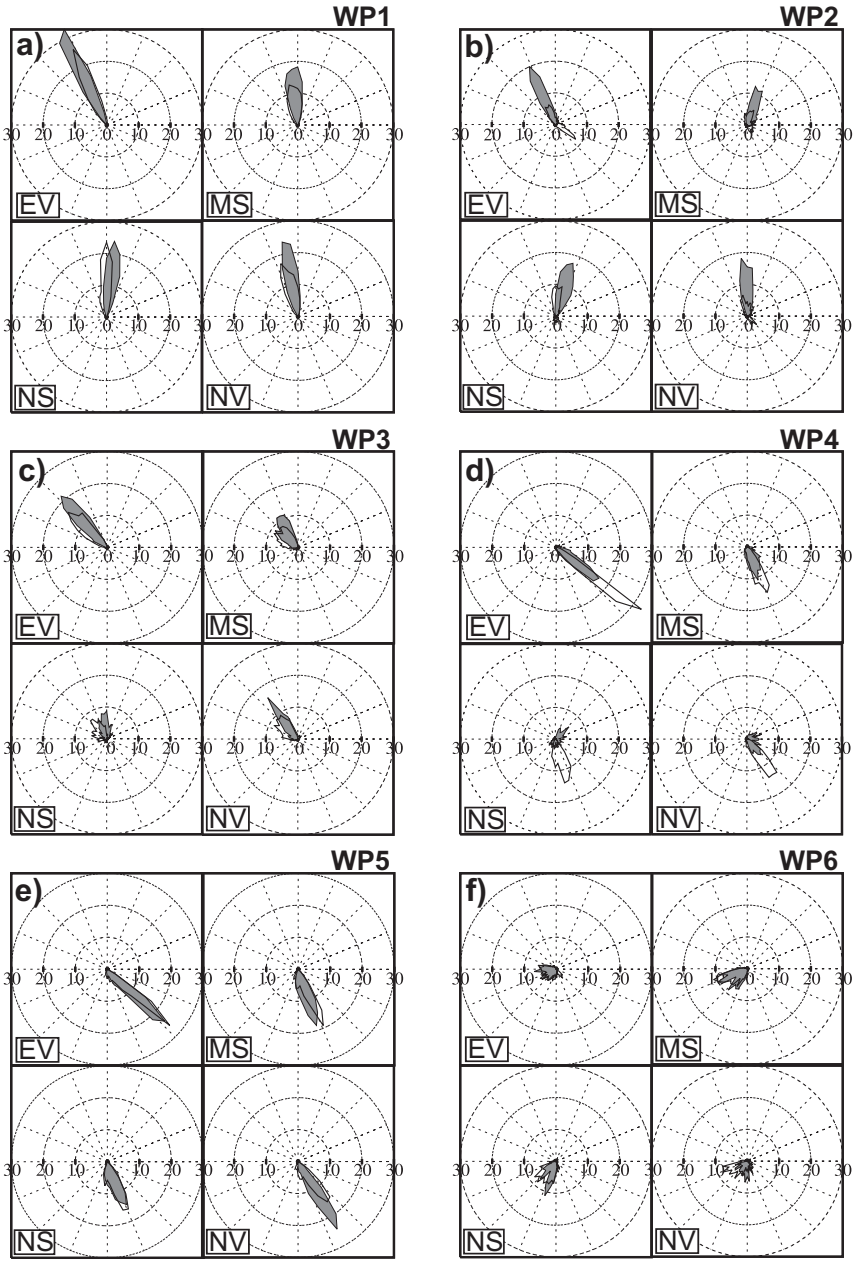

Fig. 7.6: Wind roses calculated with the observed (gray) and simulated (white) regional time series of days classified as a) WP1,b) WP2, c) WP3, d) WP4, e) WP5 and f) WP6. The homogeneous wind subregions used to calculate the regional time series are defined in Figure 4.11a.

at the four subregions (Fig. 7.6). The simulated wind roses tend to show a principal direction in agreement with those displayed by the observations. As it was expected, the main discrepancies are found for WP2, which shows two main directions in the simulated wind roses, one of them opposite to the observed one (Fig. 7.6b). The reproduction of a surface flow with an opposite sense to that of the observed case may not be limited to the poor representation of some topographical features. Instead, it appears more likely to be related to a misrepresentation of the large scale field. This will be further analyzed in the next section. The simulated wind roses at the NS and NV subregions under WP4 also show discrepancies with the observed wind roses (Figs. 7.6d). The simulation shows a coherent spatial structure with southeastern circulations in opposition to the less clear structure obtained with observations.

7.4 Large scale influence on the downscaled wind field

Previous studies have mentioned the existence of correlation between the pressure difference along the Ebro valley and the northwestern wind speed at a central location in the valley (García, 1985). The intensification of the northwestern and southeastern circulations over the CFN by this mechanism was pointed out in Chapter 6. The relationship between wind speed and the pressure gradient can be used to look for possible explanations of the discrepancies found in the wind field reproduction (Figs. 7.4, 7.5 and 7.6).

The pressure differences along the Ebro valley are calculated from daily mean observations of pressure at Santander, located at the headboard of the Ebro valley (Fig. 1.2), and at Tortosa, located at the valley mouth (Fig. 1.2). The pressure difference is defined as the observations at Tortosa minus the ones at Santander. For the simulation, this difference is calculated with the pressure at the grid points co-located to Santander and Tortosa, both selected from the outermost domain 1 (Fig. 3.1). As an illustration of the wind speed intensification exerted by the pressure gradient, the dispersion diagram between the observed (simulated) wind speed at the EV subregion and the observed (simulated) pressure along the Ebro valley is displayed in Figure 7.7a. There is a direct relationship between the mean wind speed and the absolute value of the pressure difference. The dispersion diagrams for days classified as WP1 (Fig. 7.7b) and WP5 (Fig. 7.7c) which display strong northwestern and southeastern circulations (Figs. 7.4a,e), reveal that the positive pressure differences intensify the winds down the valley (northwestern) whereas the negative differences intensify the wind up the valley (southeastern). The simulation reproduces the wind speed intensification associated with the pressure gradient although it tends to show higher wind speeds than observations for a given of the pressure gradient.

The correlation between the pressure differences along the Ebro valley and the wind speed at every subregion for days belonging to each WP are displayed for observations and simulation in Figure 7.8. The positive correlation between the pressure difference and the regional wind speeds for the northwestern WPs (i.e. WP1-3) is reproduced by the downscaling. The correlation displayed by the simulation shows a certain degradation with respect to observations for WP2, which suggests certain limitations of the downscaling to reproduce the relationship between the large scale structure and the surface

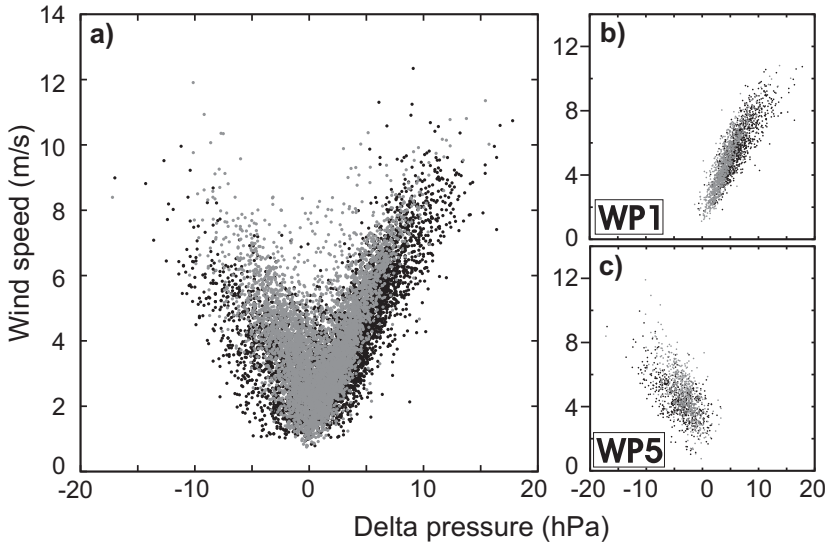

Fig. 7.7: Scattering diagram of the pressure differences along the Ebro valley and the mean wind speed field calculated with observations (black) and simulations (gray): a) for the complete dataset, b) for days classified as WP1 and c) for days classified as WP5. The pressure difference is defined as the SLP of observations/simulations at Santander minus their corresponding values at Tortosa (Fig. 1.2).

flows for days classified under this pattern. The southeastern circulations of WP4 are uncorrelated with the pressure difference in the observations, but the simulation shows some degree of association which also could indicate limitations to reproduce the relationship between large scale fields and surface circulations. The negative correlations of WP5 and WP6 show a better agreement between observations and simulation.

A more in depth evaluation of the dynamical downscaling performance as a function of the large scale situation is provided in the next section. The analysis is based on the SLP classification performed in Chapter 6.

7.4.1 Downscaling accuracy under representative synoptic situations

The SLP fields over the IP were classified into eight PPs (Fig. 6.18). The PPs showed clear associations with the six WPs (Table 6.2). Hence, an evaluation of the wind field reproducibility as a function of representative large scale situations can be obtained by analyzing the accuracy of the downscaling to reproduce the WPs under their associated PPs. For this purpose, the RMSE of the direction and wind speed bias are calculated for observations and simulation of days belonging to each one of these WP-PP associations. Four pairs of RMSE and bias are calculated within each association, one for each regional time series obtained as a result of averaging the wind information of stations classified

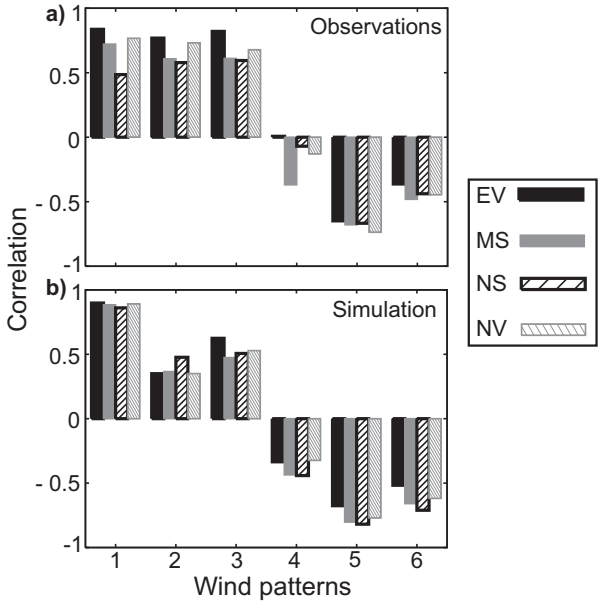

Fig. 7.8: Correlation between the regional wind speed and the pressure difference along the Ebro valley for days belonging to each WP. The upper panel displays the results obtained with observations whereas the lower panel displays the results obtained with the simulation. The homogeneous wind subregions used to calculate the regional time series are defined in Figure 4.11a.

under each wind subregion with different wind variability (Fig. 4.11a). Results are displayed in Figure 7.9. To facilitate the physical interpretation of results, the observed and simulated histograms of the pressure differences along the Ebro valley for each association between the WPs and PPs are displayed in Figure 7.10. The following paragraphs discuss the results of each WP.

The northwestern circulations of WP1 show high associations with PP1, PP4 and PP5 (Fig. 7.9a,d,e). The three PPs show higher pressures at the headboard of the Ebro valley than at its mouth which favors the intensification of the northwestern circulations of WP1. The simulation reproduces the positive pressure differences but the values are not as high as the observed ones (Fig. 7.10a,d,e). It can be argued that the underestimation of the pressure gradient along the Ebro valley contributes to the general tendency displayed by the downscaling of underestimating the wind speed at the four subregions under the three PPs (Fig. 7.9a,d,e). This follows from the positive correlation shown between these variables for days classified as WP1 (Fig. 7.8). The only exception is the NV subregion, which shows a wind speed overestimation under PP4 and PP5. The RMSE of the direction is on the range of 20 to 30 degrees under the three PPs (Fig. 7.9a,d,e), indicative of a good reproduction of the climatological wind direction of WP1. This could be a consequence

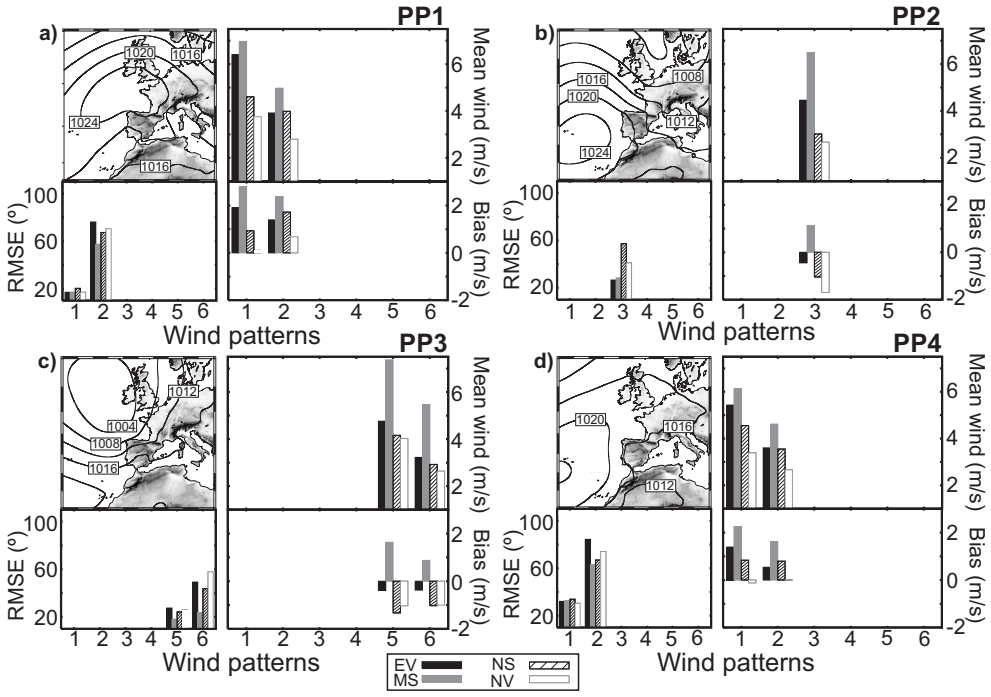

Fig. 7.9: Mean observed wind speed, wind speed bias and RMSE of the wind direction calculated with the observed and simulated regional time series of each one of the main associations encountered between the six WPs and the eight PPs (Table 6.2). The mean SLP of days belonging to each PP are also reproduced. The homogeneous wind subregions used to calculate the regional time series are defined in Figure 4.11a.

of the high wind speeds observed under this WP1 (Fig 7.5a), which minimize the effects of local orography or physical processes that are not properly modeled.

The northwestern circulations of WP2 show high associations with PP1, PP4 and PP8 (Fig. 7.9a,d,h). The three PPs show the positive pressure differences along the Ebro valley (Fig. 7.10a,d,h) that intensify the northwestern circulations over the CFN (Fig. 7.8). The simulation shows again an underestimation of these pressure differences under the three PPs. As a consequence, the simulated large scale structure favors the developing of weaker than observed wind speed circulations for days classified as WP2 (Fig. 7.9a,d,h). This wind speed underestimation is higher under PP1 than under PP4 or PP8. The RMSE of the direction shows values around 80 degrees under the three PPs. This poor performance of the numerical simulation seems to be related to the underestimation of the pressure differences along the Ebro valley previously noticed. The underestimation causes negative pressure differences in some days classified as WP2 in contrast with the positive observed ones (Fig. 7.9a,b,h). The negative differences favor

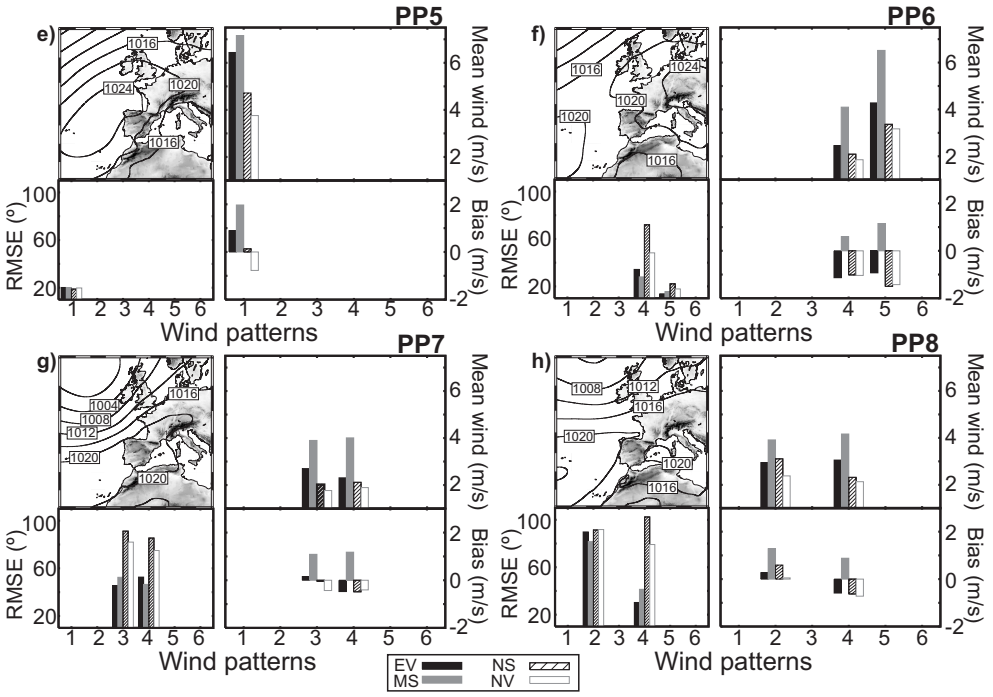

Fig. 7.9: Continuation.

the development of southeastern circulations rather than the northeastern ones. Indeed, the simulated wind display southeastern circulations in some days classified within this northwestern WP2 (Fig. 7.6b) that logically contribute to the high RMSE produced by the downscaling.

The northwestern flows of WP3 are mainly generated by PP2 (Fig. 7.9b). The wind speed is overestimated on the valley subregions (EV, NS and NV) and underestimated at the mountain subregion (MS). The wind speed overestimation occurs with a slightly underestimation of the pressure gradient (Fig. 7.10b). WP3 also shows associations with PP7 (Fig. 7.9g). The wind speed is in a better agreement with observations at the valleys but it also shows an underestimation at the MS subregion. Both PP2 and PP7 show a better reproduction of the wind direction at the EV and MS subregions than at the NV and NS subregions. The NV and NS subregions are located at areas of more complex terrain than the EV and MS subregions (Fig. 4.11a) This could be the reason of the different performance of the simulation. PP2 shows smaller RMSE than PP7. This could be a consequence of the more intense synoptic forcings in PP2 than in PP7, that produce more deterministic surface circulations and favors a better reproduction of the flow by the simulation.

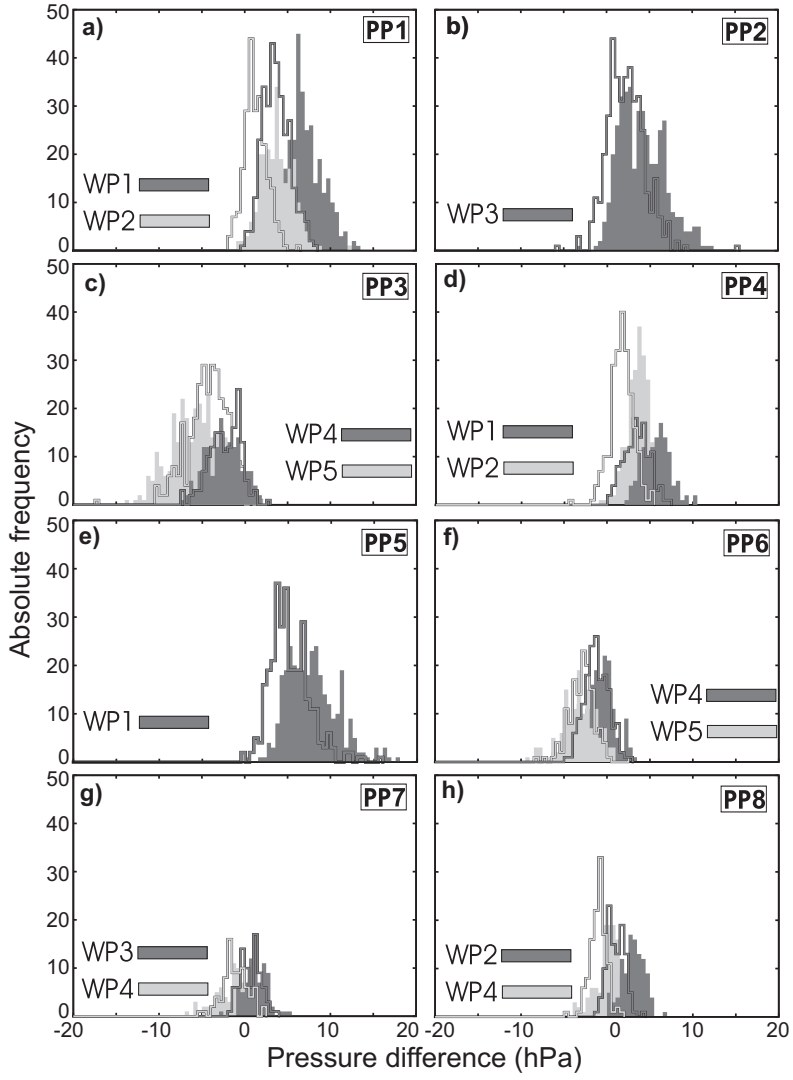

Fig. 7.10: Observed (shaded) and simulated (lines) histograms of the pressure difference along the Ebro valley for each of the main associations between the WPs and the PPs: a) PP1, b) PP2, c) PP3, d) PP4, e) PP5, f) PP6, g) PP7 and h) PP8. The pressure difference is defined as the SLP observations/simulations at Santander minus their corresponding values at Tortosa (Fig. 1.2).

The southeastern circulations of WP4 show high associations with PP6, PP7 and PP8 (Fig. 7.9f-h). The three produce a wind speed overestimation at the valleys (EV, NS and NV) and an underestimation at the mountains (MS). This WP4 does not show a relationship with the pressure gradient as high as the rest of WPs (Fig. 7.8), probably due to the weak gradients along the Ebro valley observed under the three PPs (Fig. 7.10f-h). The downscaling reproduces these near zero gradients with a slight tendency towards negative pressure differences. The negative differences tend to intensify the southeastern circulations (Fig. 7.7c), which contributes to simulate stronger southeastern circulations than the observed ones at the valleys (Fig. 7.9f-h). The RMSE of the direction is higher at the NS and NV subregions than at the EV and MS subregions for the three PPs (Fig. 7.9f-h).

The southeastern circulations of WP5 are forced by PP3 and PP6 (Fig. 7.9c,f). Under both PPs, the wind speed is overestimated at the valley subregions (EV, NS and NV) and underestimated at the MS subregion. The pressure differences along the Ebro valley matches reasonable well with observations under PP6 (Fig. 7.10f) but it is underestimated under PP3 (Fig. 7.10c). The RMSE scores are around 20 degrees for both PP3 and PP6 indicating a good reproduction of the flow direction for days classified as WP5. As WP1, this good reproduction of the wind field direction seem to be associated with the strong surface circulations typical of this pattern (Fig. 7.5a), which minimize the influence of local orographic effects or physical processes do not properly modeled.

The southwestern circulations of WP6 show important associations only with PP3 (Fig. 7.9c). As WP3, WP4 and WP5 the simulation overestimates the wind speed at the valleys and like all the previous WPs the downscaled wind is underestimated at the MS subregion. This behavior is obtained even with an appropriate representation of the pressure gradient along the valley, indicative of an appropriate reproduction of the large scale forcing (Fig. 7.10c), pointing again to the general tendency to overestimate the wind at the valleys and underestimate it at the mountains. The RMSE is around 40 degrees in the valley subregions and around 20 at the mountainous one (Fig. 7.9c).

In summary, the accuracy of the downscaling to reproduce a certain WP is somewhat independent of the large scale situation. The wind speed is generally overestimated at the valleys and it is always underestimated at the MS subregion. This seems to be related with the orographic influences discussed in section 7.3.1. Only under WP1 and WP2 the wind speed is underestimated at the valleys, but this is a consequence of limitations to reproduce the large scale which underestimate the pressure gradient along the Ebro valley and leads to weaker wind speeds over the CFN. Hence, both orographic influences and large scale misrepresentations introduce their particular biases in the downscaled wind field.

The previous evaluation analyzes the accuracy of the simulation to reproduce the climatological flow of the typical surface circulations over the CFN (WPs) and under its representative synoptic forcings (PPs). A complementary evaluation can be observed in Figure 7.11. The Taylor diagrams are calculated with the simulated and observed regional wind speed time series of days belonging to each one of the main associations among the WPs and the PPs. This comparison allows to evaluate the accuracy of the downscaling in reproducing the internal variability of the regional wind speeds within each association. The correlation is generally between 0.5 and 0.8 for all the associations. The different

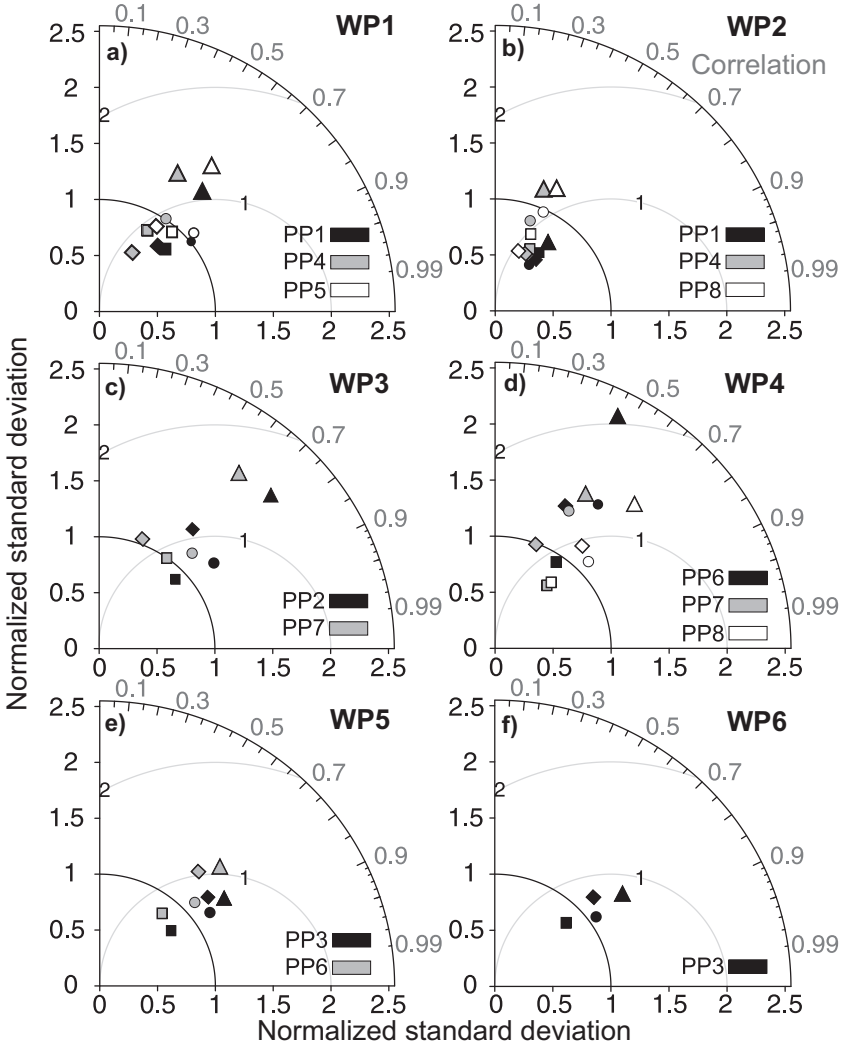

Fig. 7.11: Normalized Taylor diagrams calculated with the observed and simulated regional wind speed time series for days belonging to the main associations between the PPs and the WPs: a) WP1, b) WP2, c) WP3, d) WP4, e) WP5 and f) WP6. The symbols represent each one of the four wind subregions defined in Figure 4.11a: EV (circles), MS (squares), NS (diamonds) and NV (triangles).

subregions tend to show a similar correlation score for each association; with a lower variance than observed at the MS subregion, and higher variance than observed at the EV, NS and NV subregions in this order. For a given WP, there is not a PP under which the wind downscaling clearly outperform the rest of PPs. However, minor improvements can be observed. For instance, WP2 is better reproduced under PP1 than under PP4 or PP8; the wind speed variability of WP3 is better captured under PP2 than under PP7; WP4 is better reproduced by PP8 than by the other associations, and the same can be said for WP5 and PP3 which improve the results obtained under PP6.

The different behavior for a given WP is related with the internal variability of the wind speed for the different WP-PP associations. In order to illustrate this, Figure 7.12 displays the wind speed histograms for the EV subregion and for days classified as WP4-PP6 and WP4-PP8. The simulation shows slightly better performance in reproducing the wind speed for days classified as WP4-PP8 than WP4-PP6 (Fig. 7.11d). The histogram of the wind speed associated with PP8 is wider, higher variance, than the one of PP6. A higher variance is obviously related with a higher variability of the wind speed time series and thus, favors is better reproduction by the simulation. The other WPs that show a different behavior for different PPs support the present interpretation (not shown).

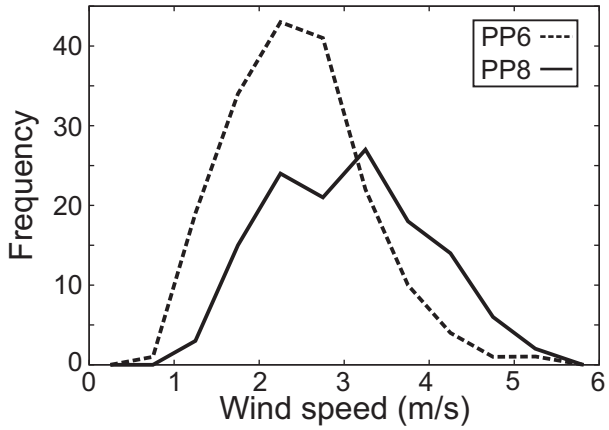

Fig. 7.12: Histograms of the wind speed observed at the EV subregion for days classified as WP4-PP6 (discontinuous line) and WP4-PP8 (continuous line).

7.4.2 Large scale misrepresentation

Two main reasons can potentially cause the pressure gradient misrepresentation: 1) deficiencies in the large scale fields from the GCM used as initial and boundary conditions by the WRF downscaling, or 2) deficiencies in the WRF simulation to reproduce the large scale fields. To analyze both possibilities it turns out necessary to analyze the capability

of both simulations to reproduce the observed SLP field. For this purpose, observational SLP fields were used (Trenberth and Paolino, 1980, source: ds010.1 of the NCAR archive). The SLP fields consist of monthly averaged values over a spatial grid of 5 degrees lat x lon. Hence, the SLP fields from ECMWF and WRF grids were monthly averaged and interpolated to a 5 degree spatial resolution in order to obtain comparable datasets.

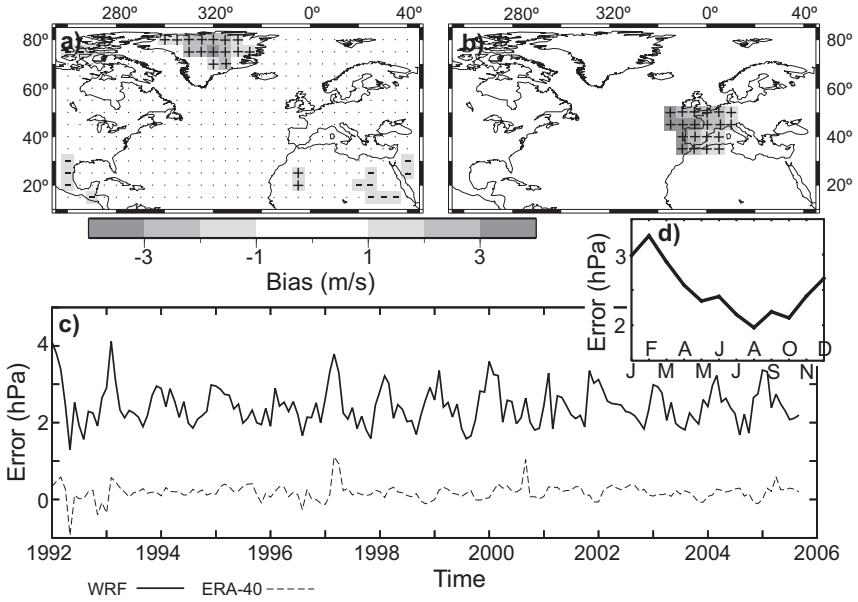

Fig. 7.13: Bias between observational SLP fields from the ds010.1 dataset of the NCAR archive and the ECMWF (a) and WRF simulations (b). The time series of the spatially averaged differences of each month are displayed in panel (c), and its annual cycle for the WRF simulations in panel (d).

The bias between the observed and simulated SLP fields from the ECMWF and WRF models are displayed in Figures 7.13. The ECMWF data shows small deviations with respect to the mean of the observed values, which are below 1 hPa for the areas surrounding the IP (Fig. 7.13a). On the contrary, the downscaled SLP fields show a general underestimation (Fig. 7.13b). This underestimation is higher at western areas than at eastern areas of the IP, leading to the underestimation of the pressure differences along the Ebro valley. The time series of the spatially averaged departures from the observed SLP field are represented for both the ECMWF and the WRF fields in Figure 7.13c. Only the grid points available to calculate the WRF time series (Fig. 7.13b) are used in the calculation of the ECMWF time series of differences. The ECMWF shows an error evolution very close to zero with a few sporadic peaks, whereas the downscaled fields present a systematic underestimation. An annual cycle with higher errors in winter

than in summer can be recognized. This is better appreciated in Figure 7.13d which shows the monthly means of the downscaled errors.

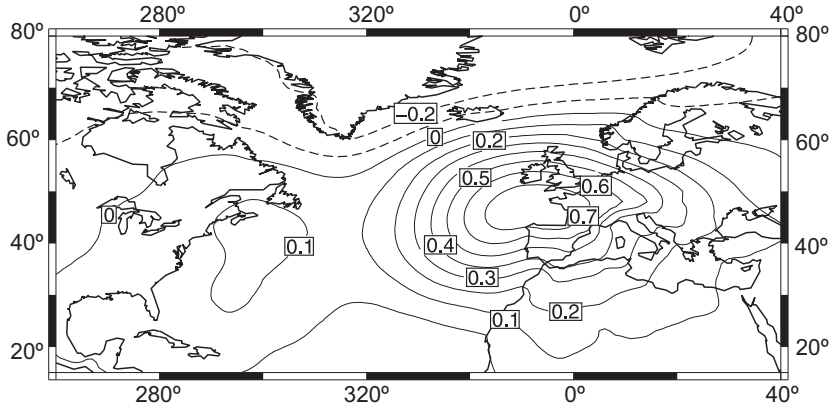

Fig. 7.14: Correlation between the anomalies of the daily SLP fields of the ECMWF datasets and the error time series of the WRF simulation displayed in Figure 7.13c. Positive (negative) values are represented in continuous (dashed) lines.

To identify the large scale situation that produces the SLP misrepresentation in the downscaling, a correlation map is calculated with the time series of the spatially averaged departures produced by the WRF downscaling (Fig. 7.13c) and the SLP fields from the ECMWF datasets. The deterministic annual cycle was removed from the time series of the departures and the SLP fields before calculating the correlation. The anomalies are calculated by subtracting to each day its monthly climatological mean. The correlation map is displayed in Figure 7.14. There is a center of high correlation ($r > 0.7$) to the northwest of the IP. Hence, the higher the pressure over this area the larger is the large scale misrepresentation. Areas located far away from the IP are uncorrelated with the time series of the departures. The correlation pattern reveals deficiencies in the downscaling approach herein adopted to capture the high pressures over the northwest of the IP, which cause an inappropriate simulation of the pressure gradient along the Ebro valley being partially responsible for the wind speed discrepancies encountered.

7.5 Conclusions

The accuracy of the WRF dynamical downscaling to reproduce the climatological surface circulations over the CFN has been analyzed. The analysis evaluates of the capability of the downscaling to reproduce the climatological flow of the six typical wind fields (or WPs) over the region identified in Chapter 6. Performing the analysis at the daily time

scales allowed to focus on dynamically driven circulations associated with the atmospheric dynamics and its interaction with the orography.

The spatial structure of the flow of each WP is generally captured by the downscaling. The wind speed is systematically underestimated at the sites located at mountain tops. This was attributed to limitations of the horizontal resolution used in the simulation, 2 km, to reproduce the actual height of the mountains. On the valleys, the downscaling shows a tendency to overestimate the wind speed probably due to the drag exerted by the subgrid-scale orography which is not parametrized in the present simulation. The underestimation at the windiest locations (mountain tops), and the overestimation at the less windy sites (valleys) reduces the spatial variability of the downscaled wind speed field in comparison with the observed one.

Contrary to the general behavior of the downscaled wind, the northwestern WP1 and WP2 show an underestimation of the wind speed at the valleys. This is a consequence of the underestimation of the SLP to the north of the IP, which produces weaker pressure gradients along the Ebro valley than the observed ones, and ultimately lead to the underestimation of the northwestern circulations of these WPs. The SLP underestimation produces in some days classified as WP2 an opposite pressure gradient along the Ebro valley than the observed one, which led to the simulation of southeastern circulations instead of the northwestern flows typical of WP2. These results show the strong influence that the large scale exerts over the surface circulations, and suggests that the potential improvement obtained in the downscaling by increasing the horizontal resolution may be hampered if the large scale is not appropriately captured.

8

Conclusions and something else

Perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances and at a cost less than the saving to mankind due to the information gained. But that is a dream.

L. F. Richardson, 1922:
Weather Prediction by Numerical Process. Cambridge University Press.

The surface circulations over the CFN have been analyzed in observations and in a numerical simulation performed with the WRF model (Skamarock et al., 2005). The complete observational period from 1 January 1992 to 7 October 2005 was simulated at a high horizontal resolution over the CFN (2 km) to allow for an accurate evaluation of the simulation performance in reproducing the surface wind behavior. It should be emphasized the innovative character of this simulation in its length and spatial resolution.

The analysis focused on daily means to filter out thermally induced circulations and concentrate the study on dynamically induced circulations exerted by the large scale flow and its interaction with the topographic features.

The following section summarizes the specific conclusions of each chapter of the Thesis.

8.1 Conclusions

The original observations were subjected to a quality control and a bias correction to suppress potentially erroneous records and confer reliability to the conclusions reached in the analysis.

- About 1.8 % wind speed and 3.7 % wind direction records were diagnosed as invalid. The high percentage of records invalidated stresses the importance of this kind of depuration process in order to reach solid conclusions in the analysis. In spite of the beneficial effects of the QC and the correction of biases, there is no guarantee that other kind of erroneous records still persists in the dataset. These undetected errors can contribute to the discrepancies encountered in the comparison of the observations

with the WRF simulation. Nevertheless, the QC applied to the data ensures that the dataset has improved its quality with respect to its initial state.

The wind variability was analyzed by classifying together those observational sites with similar temporal variability by performing a wind regionalization on the observational and simulated datasets.

- A total of four wind regions were identified showing a distribution in accordance with topographic features of the terrain. The meridional wind variability is rather similar over the four subregions indicating an homogeneous behavior of this wind component over the whole CFN. The zonal wind variability is therefore responsible for the different wind behavior on the subregions.
- Spectral and wavelet analysis revealed the presence of a dominant annual wave and varying presence of higher frequency contributions. The valley subregions tend to show more activity at higher frequencies than the mountain subregion.
- The evaluation of the simulation at the regional scale has been shown to provide advantages in comparison with the evaluation at the sites. The averaging of the individual time series of each region reinforces the regional signal and mitigates local effects in observations and *representativeness errors* in the simulation providing more appropriate time series for comparison.
- The WRF simulation reproduces reasonable well the meridional wind variability over the CFN, and the zonal wind variability at most of the identified subregions. The Ebro valley and the mountain stations are the subregions wherein the simulation displayed a better performance in reproducing the wind variability.
- The higher spatial and temporal coverage of the simulation was used to infer the wind variability at areas and periods with scarcity of observations. For this purpose, the regionalization methodologies were applied to the simulated wind. The identified subregions are consistent with those obtained with observations over the CFN, but the higher coverage of the simulation reveals the existence of additional subregions outside of the network coverage. These additional subregions appear well defined covering large areas around the CFN, and suggest the use of observations over these areas on future studies involving the wind variability over the CFN.
- The inference analysis provides valuable information for the design or improvement of observational networks. It allows to distribute the stations in order to appropriately capture the wind variability over the area. More density of stations in the smaller regions would be desirable. This indicates that a homogeneous distribution of stations could not be the best distribution of stations.

The climatological circulations over the area were analyzed by classifying together those daily wind fields with similar surface flows into typical patterns or WPs.

- A total of six WPs were identified. The Northwestern circulations were found to be dominant (60.9 %) followed by the southeastern ones (30.5 %). This is the main direction of the valleys over the region and shows the strong influence that the orography produces over the surface circulations.

- In order to understand their forcing mechanisms, the SLP fields over the IP for the 1992 to 2005 period were classified into typical pressure patterns or PPs, and related to the WPs already identified. The relationships found between the PPs and the WPs revealed the importance of the ageostrophic balance and the intensification of the flow due to the pressure gradient difference between the headboard and the mouth of the Ebro valley.
- The climatological wind direction was better reproduced on the mountain subregion that at the valley subregions likely due to the less influence that the orography produces on the former locations. The simulation tends to overestimate the climatological mean wind speed at the valley subregions, and to underestimate it at the mountain subregion. The underestimation of the wind speed at the windiest sites (mountain sites) and the overestimation at the less windy sites (valleys) leads to an underestimation of the spatial wind speed variability. It was shown that this behavior is at least partially related with the smoother topography used in the simulation.
- In concordance with the long term climatology explained above, the simulation shows systematic wind speed underestimation at the mountain subregion under the six WPs. Also in agreement with the general behavior was the overestimation of the wind speed over the valleys under four WPs; however, two northwestern WPs showed a wind speed underestimation. This was a consequence of a large scale misrepresentation in the downscaling.
- Both, limitations to represent the terrain features and limitations to reproduce the large scale circulations introduce errors in the surface wind estimations. Hence, the expected improvement as a result of increasing the horizontal resolution may be at least partially hampered if the large scale is not appropriately reproduced. Further research turns necessary to understand the causes that generated the misrepresentation of the large scale structure. A way around to circumvent the problem can be attained by applying nudging techniques to the outermost domain in order to impose the large scale structure to the downscaling.

8.2 Quo Vadis?

The dynamical downscaling has shown a reasonable good accuracy in reproducing the wind variability induced by dynamical forcings over the area, with certain biases in reproducing the climatological flow. This finding opens new research questions that can be addressed in future studies. The following sections provide details for some of these potential lines.

Extension of the simulation

The good performance of the simulation in reproducing the wind variability over the CFN is summarized in Figure 8.1 that shows the observed and simulated regional wind components. The variations in the simulated and observed time series are in quite good agreement ($r = 0.87$ and $r = 0.93$ for the zonal and meridional components, respectively), which encourages to extend the simulation in order to analyze the wind variability before

1992, when observations are very limited. The enlargement of the simulation is currently under development and at the time of editing this text the simulation has been extended back to 1960. The simulated meridional wind component shows the already noticed tendency of the downscaling to overestimate the southeastern wind circulations (positive values) and to underestimate the northwestern circulations (negative values).

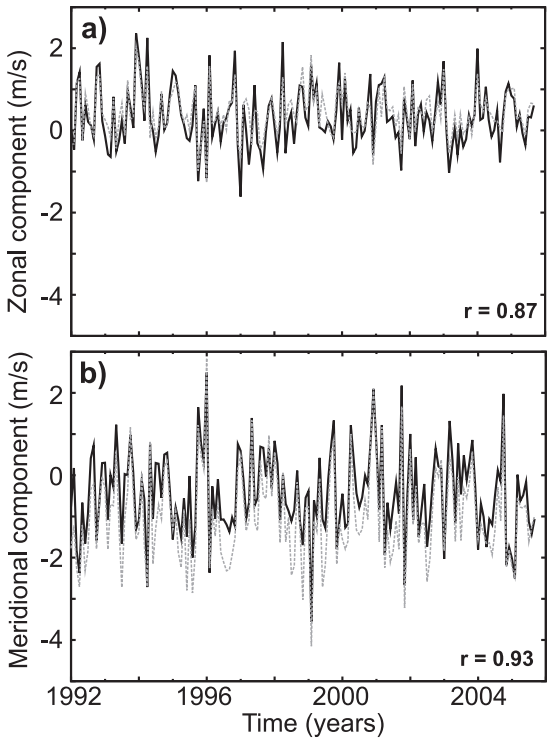

Fig. 8.1: Monthly mean of the spatially averaged zonal (a) and meridional (b) wind components at the 41 observational sites (Fig. 1.2). The black (gray) lines represent the simulation (observations).

The use of the extended simulation to analyze the climatological flow behavior should be therefore taken with care. A certain correction seems to be necessary to remove the systematic errors introduced by the downscaling. One possible correction could be based on the use of the SLP classification (Chapter 6) to calculate the biases introduced on the surface wind estimations under each PP. Then, the simulated wind during a certain temporal period, let's say a month, can be corrected by adding the bias of each PP weighted by their relative frequency of appearance on that month. The effects of this correction on the estimations of the regional wind components are displayed in Figure 8.2. The corrected time series are in a better agreement with observations than the

direct outputs of the downscaling (Fig. 8.1), specially the ones of the meridional wind component. Still, some discrepancies are evidenced which can potentially be improved in the future by applying more sophisticated corrections. The longer time series provided by the simulation allow for an analysis of the wind behavior since 1960, a period much longer than the one covered by observations. For instance, a first look reveals that there were not large changes in the low frequency variations during the last decades.

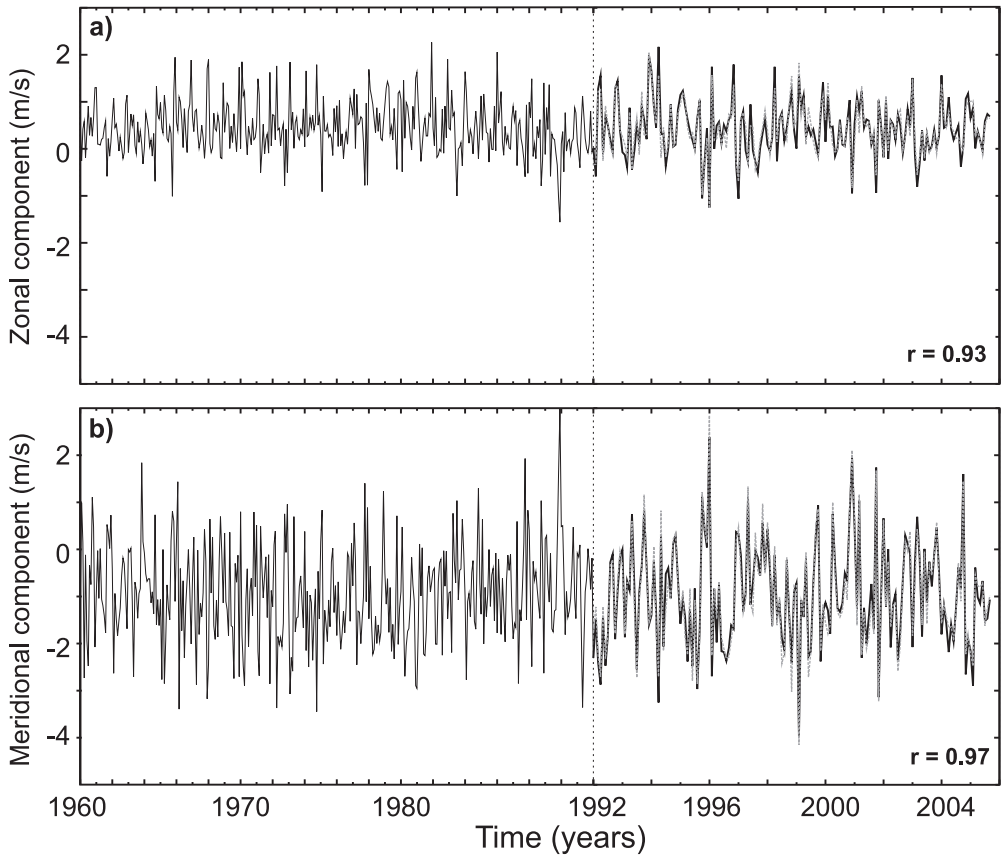

Fig. 8.2: Monthly mean of the spatially averaged zonal (a) and meridional (b) wind components at the 41 observational sites (Fig. 1.2). The black lines represent the simulation and the gray lines the observations.

Testing of different physical packages

The knowledge of the different regions of wind behavior over the CFN (the regionalization), the climatological surface circulations over the region (the WPs) as well as their relationship with the large scale situations (PPs) convert the CFN into and ideal benchmark to test downscaling sensitivities to the mesoscale model configuration. For instance, the influence on the estimated wind field as a result of changing the model physics can be evaluated. In order to illustrate the potential application of the classifications herein performed, a complete year was simulated with the MYJ PBL parameterization (Mellor and Yamada, 1982; Janjić, 1994) and the outputs compared with those obtained with the YSU package. The distribution of the mean wind speed field differences for days belonging to each WP are displayed in Figure 8.3. The estimations using the MYJ package produce higher wind speeds than the estimations that use the YSU physics under the six WPs. This overestimation helps to produce wind speeds in better agreement with observations in WP1 and WP2 whose mean wind speed was underestimated by the YSU package. However, it produces a negative influence on WP3-6 under which the wind speed field was overestimated by the estimations performed using YSU. It would be interesnting to analyze the different contributions that the YSU and MYJ packages produce on the simulated WP during the night and day. The comparison can be enlarged by analyzing the performance of the estimations on each wind region or under each PP. These comparisons can help to understand the influence over the surface circulations of the different PBL physic schemes and perhaps to improve their performance.

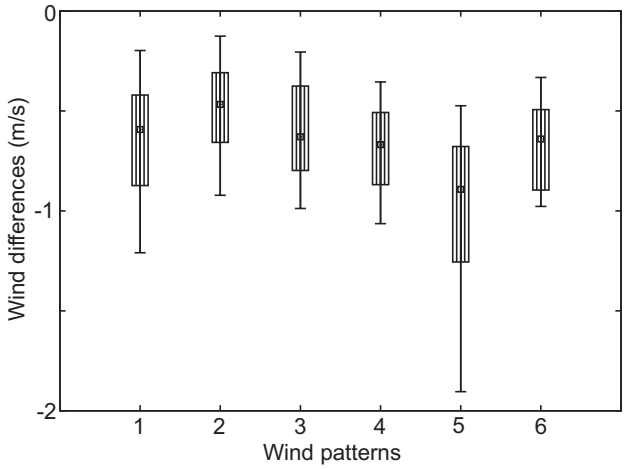

Fig. 8.3: Box and whiskers plots showing the distribution of the differences of the mean wind speed fields obtained with the YSU and ETA PBL scheme for each WP. The dot indicates the median, the lower bases of the boxes are the 25 percentiles and the upper ones the 75 percentiles. The lower and upper whiskers indicate the 10 (lower) and 90 (upper) percentiles.

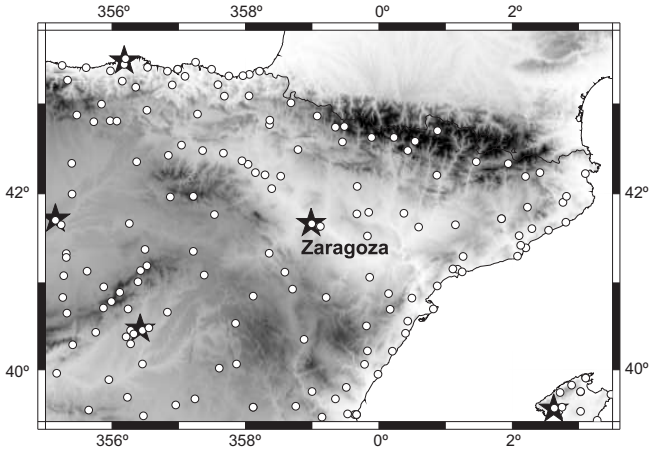

Fig. 8.4: Ebro valley basin topography (shaded) and locations of the surface stations with wind observations available from AEMET, and locations with radiosonde data (stars) from the Integrated Global Radiosonde Archive (IGRA, Durre et al., 2006; Durre and Yin, 2008).

Extension of the spatial coverage: including wind observations at upper levels.

The strong influence that the presence of the Ebro valley exerts over the CFN surface circulations has been shown within this Thesis. It would be therefore interesting to enlarge the area of study in the future to completely include the Ebro basin. There is a considerable number of surface observations available from the Spanish meteorological agency (AEMET) to provide a reasonable good spatial coverage of the surface circulations within the valley (Fig. 8.4). In addition, there are other mesonets, such as the one from the CFN used herein, that will allow to increase the spatial resolution in the surface wind sampling. The extension of the area of study would not only allow to analyze the surface circulations along the complete basin, but it will also allow to include in the area of study locations with available radiosonde data (Fig. 8.4). For instance, the radiosonde at Zaragoza, located in the middle of the Ebro valley, can be used to analyze the vertical structure of the surface circulations along the valley. The wind speed vertical profile as a result of averaging those days classified as WP1, strong circulations down the valley, and days classified as WP5, strong circulations up the valley, can be observed in Figure 8.5a and b, respectively. The northwestern circulations, donwnvalley, show wind speed maxima at lower levels (Fig. 8.5a) that do not appear in the southeastern circulations, upvalley (Fig. 8.5b). It would be interesting to analyze in future research the causes of this observed behavior.

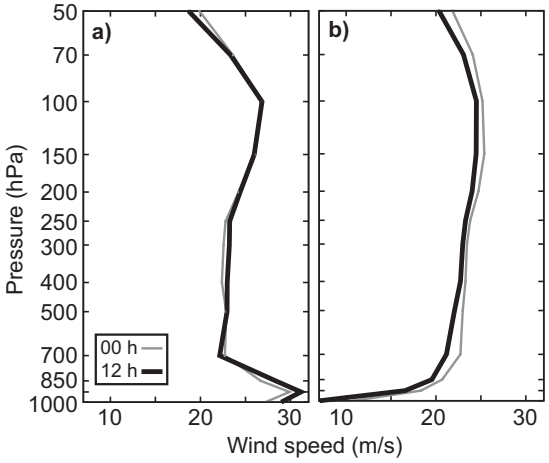

Fig. 8.5: Wind speed vertical profile for days classified as a) WP1 and b) WP5 at Zaragoza (8.4). The gray (black) line is associated with the radiosonde taken at 00 (12) hours.

Intradiurnal variability: on the effects of thermal forcings

The thesis focused on the wind variability and climatological flow behavior at daily time scales. The circulations at this temporal scale are mainly related to dynamically driven circulations associated with large scale atmospheric variations. Considering sub-daily time scales would allow for the analysis of thermally driven circulations associated with the diurnal solar cycle. The heating and cooling of the ground interacts with the nearby air producing a variety of physical processes that generate diurnal variations on the atmospheric flow. A first step in evaluating the influence of thermal effects over the CFN surface circulations could be accomplished by analyzing the climatological diurnal wind variations (Jiménez et al., 2009e). The analysis should be undertaken with summer days due to the stronger diurnal wind variations over land areas outside the tropics, like the CFN, during this season (Dai and Desser, 1999). As a preliminary overview, the climatological wind speed diurnal cycles measured at the 41 stations during summer are displayed in Figure 8.6. It can be appreciated that large variations exist in mean velocity between the different locations, which range from less than 1 ms^{-1} to more than 10 ms^{-1}. The wind speed values are concentrated in a narrower interval during central hours of the day, since locations with high wind speed tend to present a minimum, and locations with low wind speeds a maximum. The stations with a minimum at central hours of the day tend to be locations well exposed to the stronger general winds like the mountain tops or hills in the Ebro valley. The different behavior produces a tendency to homogenize the wind speed over the area during the day, and a more heterogeneous behavior during the night. It seems to be a consequence of the turbulent mixing that tends to homogenize the properties within the PBL. This mixing is stronger at central hours of the day, just when the more homogeneous wind is observed. During the night, the turbulent mixing

decreases its intensity and allows for an stratification of the meteorological variables with height and produces the more heterogeneous wind speed over the area. The validity of this hypothesis can be tested using the WRF simulation as an approximation of the real world. Another feature that can be recognized in the diurnal variations (Fig. 8.6) is a general tendency to present a wind speed maximum (minimum) at later hours of the day (night). It seems to be associated with the effects of the sea-breeze circulations over the area (Jiménez et al., 2009e), which favors (is opposite to) the mean surface wind during the day (night).

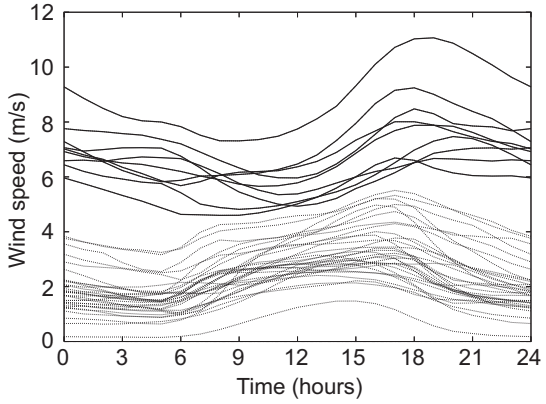

Fig. 8.6: Diurnal wind speed variations as a result of averaging the summer observations at each one of the 41 observational stations (Fig. 1.2). The thin (thick) lines lines highlight the sites with a wind speed maxima (minima) at central hours of the day.

Sensitivities to soil characteristics

Another interesting analysis for future research would be the use the WRF model to test sensitivities of the surface circulations to certain physical properties such as the surface characteristics. Among the different surface physical properties, soil moisture availability seems to be an interesting candidate since it plays a major role in the land-atmosphere interactions (Seneviratne et al., 2006). An example of the effects introduced on the diurnal wind speed variations during summer at a central location in the CFN as a result of changing the surface moisture availability is displayed in Figure 8.7. A drier soil favors the development of higher wind speeds. The largest differences appear at the late afternoon, that shows wind speed values that differ up to 2 ms^{-1} between the different simulations. The causes that origin these differences are currently under study. The roughness length is another interesting surface property to test, since it plays an important role in the vertical wind speed profile (Stull, 1988).

In short, the study herein described opens an interesting range of topics for future research which will contribute to increase our understanding of the surface wind behavior

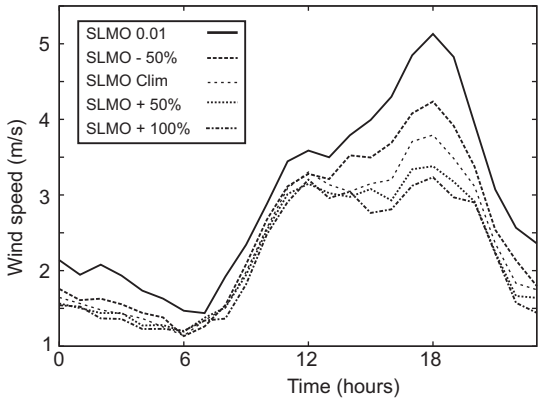

Fig. 8.7: Diurnal wind speed variations as a result of averaging simulations performed with different values of the soil moisture availability for the summer 2003.

as well as the performance of numerical simulations in reproducing it. There is still considerable research that needs to be undertaken to use the numerical simulations as a substitute for observations, but there is little doubt about the fulfillment of Richardson's (1922) dream at the beginning of the XX century to approximate a solution of the atmospheric equations of motion. Perhaps, a new challenge would be to find the analytical solution of the atmospheric equations of motion, if there is any, and use it to calculate an ensemble of millions of solutions departing from closer atmospheric states to study the chaotic nature of the atmosphere, but of course, that's a dream.

References

Aguilar, E., I. Auer, M. Brunet, T. C. Peterson, and J. Weringa, 2003: Guidelines on Climate Metadata and Homogenization. Technical Report WMP-TD no. 1186, WMO, Geneva.

Alexanderson, H. A., 1986: A homogeneity test applied to precipitation data. *J. Climatol.*, **6**, 661–675.

Anderberg, M. R., 1973: *Cluster Analysis for Applications*. Academic Press, 359 pp.

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer, 1976: A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964.

Arakawa, A. and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Meth. Comput. Phys*, **12**, 173–265.

Archer, C. L. and M. Z. Jacobson, 2003: Spatial and temporal distributions of U.S. winds and wind power at 80 m derived from measurements. *J. Geophys. Res.*, **108(D9)**, 4289, doi:10.1029/2002JD002076.

— 2004: Correction to "Spatial and temporal distributions of U.S. winds and wind power at 80 m derived from measurements". *J. Geophys. Res.*, **109**, D20116, doi:10.1029/2004JD005099.

— 2005: Evaluation of global wind power. *J. Geophys. Res.*, **110**, D12110, doi:10.1029/2004JD005462.

Bailey, B. H. and S. L. McDonald, 1997: *Wind resource assessment handbook*. AWS Scientific, Inc, 79 pp.

Baker, R. W., E. W. Hewson, N. G. Butler, and E. J. Warchol, 1978: Wind power potential in Pacific Northwest. *J. Appl. Meteor.*, **17**, 1814–1826.

Bakun, A. and V. N. Agostini, 2001: Seasonal patterns of wind-induced upwelling/downwelling in the Mediterranean Sea. *Scientia Marina*, **65**, 243–257.

Banta, R. M., L. S. Darby, P. Kaufmann, D. H.Levinson, and C.-J. Zhu, 1999: Wind-flow patterns in the grand canyon as revealed by doppler lidar. *J. Appl. Meteor.*, **38**, 1069–1083.

Barnett, T. P., 1977: The principal time and space scales of the Pacific trade wind field. *J. Atmos. Sci.*, **34**, 221–236.

Bärring, L., 1988: Regionalization of daily rainfall in Kenya by means of common factor analysis. *J. Climatol.*, **8**, 371–389.

Beljaars, A. C. M., 2004: A new parametrization of turbulent orographic form drag. *Quarter. J. Roy. Met. Soc.*, **130**, 1327–1347.

Belserene, E. P., 1988: Rhythms of a variable star. *Sky and Telescope*, **76**, 288–290.

Berg, L. K. and S. Y. Zhong, 2005: Sensitivity of MM5-simulated boundary layer characteristics in turbulence parameterizations. *J. Appl. Meteor.*, **44**, 1467–1483.

Biel, A. and L. García, 1962: El clima de Zaragoza y Ensayo Climatológico para el Valle del Ebro. Technical Report Serie A, no. 36, INM, Madrid.

Bjerknes, V., 1904: Das problem der wettervorhersage, betrachtet vom standpunkte der mechanik und der physik. *Meteorol. Z.*, **21**, 1–7.

Black, T. L., 1994: The new NMC mesoscale Eta Model: Description and forecast examples. *Wea. Forecasting*, **9**, 265–278.

Bliss, N. B. and L. M. Olsen, 1996: Development of a 30-arc-second digital elevation model of South America. *In: Pecora Thirteen, Human interactions with the Environment-Perspectives from Space*, Sioux Falls, South Dakota, August 20-22.

Bloomfield, P., 1976: *Fourier analysis of time series: An Introduction*. New York: John Wiley and Sons, 258 pp.

Bogardi, I. and I. Matyasovszky, 1996: Estimating daily wind speed under climate change. *Solar Energy*, **57**, 239–248.

Bonell, M. and G. Sumner, 1992: Autumn and winter daily precipitation areas in Walles, 1982-1983 to 1986-1987. *Int. J. Climatol.*, **12**, 77–102.

Bonnardot, V. and S. Cautenet, 2009: Mesoscale atmospheric modeling using a high horizontal grid resolution over complex coastal terrain and a wine region of South Africa. *J. Appl. Meteor. & Climatol.*, **48**, 330–331.

Brücher, W., M. J. Kerschgens, and F. Steffany, 1994: On the generation of synthetic wind roses in orographically structured terrain. *Theor. Appl.Climatol.*, **48**, 203–207.

Brunet, M., O. Saladie, P. Jones, J. Sigro, E. Aguilar, A. Moberg, D. Lister, A. Walther, D. Lopez, and C. Almarza, 2006: The development of a new dataset of Spanish daily adjusted temperature (sdats) (1850-2003). *Int. J. Climatol.*, **26**, 1777–1802.

Bubnova, R., G. Hello, P. Bernard, and J.-F. Geleyn, 1995: Integration of the fully elastic equations cast in the hydrostatic pressure terrain following coordinate in the framework of the ARPEGE/ALADIN NWP system. *Mon. Wea. Rev.*, **123**, 515–535.

Buckley, R. L., 2004: Statistical comparison of Regional Atmospheric Modelling System forecasts with observations. *Meteorol. Appl.*, **11**, 67–82.

Bunkers, M. J. and J. R. Miller Jr., 1996: Definition of climate regions in the Northern plains using an objective cluster modification technique. *J. Climate*, **9**, 130–146.

Burlando, M., M. Antonelli, and C. F. Ratto, 2008: Mesoscale wind climate analysis: identification of anemological regions and wind regimes. *Int. J. Climatol.*, **28**, 629–641.

Busuioc, A., C. Deliang, and C. Hellström, 2008: Performance of statistical downscaling models in GCM validation and regional climate change stimates: application for for Swedish precipitation. *Int. J. Climatol.*, **21**, 557–578.

Cahalan, R. F., L. Oreopoulos, A. Marshak, K. F. Evans, A. B. Davis, R. Pincus, K. H. Yetzer, B. Mayer, R. Davies, T. P. Ackerman, H. W. Barker, E. E. Clothiaux, R. G.

Ellingson, M. J. Gray, E. K. Assianov, S. Kinne, A. Macke, W. O'Hirok, P. T. P. andS. M. Prigarin, A. N. Rublev, G. L. Stephens, F. Szczap, E. E. Takar, T. Várnai, G. Wen, and T. B. Zhuravleva, 2005: The I3RC: bringing together the most advanced radiative transfer tools for cloudy atmospheres. *Bull. of the Amer. Met. Soc.*, **86**, 1275–1293.

Cairns, M. M. and J. Corey, 2003: Mesoscale model simulations of high-wind events in the complex terrain of western Nevada. *Wea. Forecasting*, **18**, 249–263.

Cattell, R. B., 1966: The scree test for the number of factors. *Multivar. Behav. Res.*, **1**, 245–276.

Chandrasekar, A., C. R. Philbrick, R. Clark, B. Doddridge, and P. Georgopoulos, 2003: Evaluating the performance of a computationally efficient MM5/CALMET system for developing wind field inputs to air quality models. *Atmos. Environ.*, **37**, 3267–3276.

Charney, J. G., R. Fjortoft, and J. Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237–254.

Chatfield, C. and A. J. Collins, 1980: *Introduction to multivariate analysis*. Chapman and Hall, 246 pp.

Cheng, C. S., G. Li, Q. li, and H. Auld, 2007: Statistical downscaling of hourly and daily climate scenarios for various meteorological variables in South-central Canada. *Theor. Appl.Climatol.*, **91**, 129–147.

Cheng, E. D., 1998: Macroscopic extreme wind regionalization. *J. Wind Eng. Ind. Aerodyn.*, **77&78**, 13–21.

Christensen, O. B., J. H. Christensen, B. Machenhauer, and M. Botzet, 1998: Very high resolution regional climate simulations over Scandinavia. *J. Climate*, **12**, 3204–3229.

Clarkson, D. B. and R. I. Jennrich, 1988: Quartic rotation criteria and algorithms. *Psychometrika*, **53**, 251–259.

Comrie, A. C. and E. C. Glenn, 1998: Principal components-based regionalization of precipitation regimes across the southwest United States and northern Mexico, with an application to monsoon precipitation variability. *Clim. Res.*, **10**, 201–215.

Conil, S. and A. Hall, 2006: Local regimes of atmospheric variability: A case study of southern California. *J. Climate*, **19**, 4308–4325.

Coronato, F. and A. Bisigato, 1998: A temperature pattern classification in Patagonia. *Int. J. Climatol.*, **18**, 765–773.

Cotton, W. R., R. A. Pielke, R. L. Walko, G. E. Liston, C. Tremback, H. Jiang, R. L. McAnelly, J. Y. Harrington, M. E. Nicholls, G. G. Carrion, and J. P. McFadden, 2003: RAMS 2001: Current status and future directions. *Meteorol. Z.*, **82**, 5–29.

Cox, R., B. L. Bauer, and T. Smith, 1998: A mesoscale model intercomparison. *Bull. of the Amer. Met. Soc.*, **79**, 265–283.

Cushman-Roisin, B., 1994: *Introduction to geophysical fluid dynamics*. Prentice Hall, 320 pp.

Dai, A. and C. Desser, 1999: Diurnal and semidiurnal variations in global surface wind and divergence fields. *J. Geophys. Res.*, **104**, 109–125.

Darby, L. S., 2005: Cluster Analysis of Surface Winds in Houston, Texas, and the Impact of Wind Patterns on Ozone. *J. Appl. Meteor.*, **44**, 1788–1806.

Davis, R. E., B. P. Hayden, D. A. Gay, W. L. Phillips, and G. V. Jones, 1997: The north atlantic subtropical anticyclone. *J. Climate*, **10**, 728–744.

Davis, R. E. and L. S. Kalkstein, 1990: Development of an automated spatial synoptic climatological classification. *Int. J. Climatol.*, **10**, 769–794.

Debernard, J., Ø. Sætra, and L. P. Røed, 2002: Future wind, wave and storm surge climate in the northern North Atlantic. *Clim. Res.*, **23**, 39–49.

Debernard, J. B. and L. P. Røed, 2008: Future wind, wave and storm surge climate in the Northern Seas: a revisit. *Tellus*, **60A**, 427–438.

Deeming, T. J., 1975: Fourier analysis with unequally-spaced data. *Astr. Spa. Sci.*, **36**, 137–158.

Defant, F., 1949: Zur theorie der Hangwinde, nebst Bemerkungen zur Theorie der Beg- und Talwinde. *Arch. Meteor. Geophys. Bioklim.*, **A1**, 421–450.

DeGaetano, A. T., 1996: Delineation of mesoscale climate zones in the Northeastern United States using a novel approach to cluster analysis. *J. Climate*, **9**, 1765–1782.

— 1997: A quality-control routine for hourly wind observations. *J. Atmos. Oceanic Technol.*, **14**, 308–317.

— 1998: Identification and Implications of Biased in U.S. Surface Wind Observation, Archival, and Summarization Methods. *Theor. Appl.Climatol.*, **60**, 151–162.

Della-Marta, P. M., 2006: A method for hogenizing the extremes and mean of daily temperature measurements. *J. Climate*, **19**, 4179–4197.

Dickinson, R. E., R. M. Errico, F. Giorgi, and G. T. Bates, 1989: A regional climate model for the western United States. *Climatic Change*, **15**, 383–422.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107.

— 1996: A multilayer soil temperature model for MM5. *Preprints, Sixth PSU/NCAR Mesoscale Model Users' Workshop*, Boulder, CO 80307, 49–50.

Dudhia, J., D. Gill, K. Manning, W. Wang, and C. Bruyere, 2004: *PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modelling System Version 3*.

Durre, I., M. J. Menne, and R. S. Vose, 2008: Strategies for evaluating quality assurance procedures. *J. Appl. Meteor. & Climatol.*, **47**, 1785–1791.

Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the integrated global radiosonde archive. *J. Climate*, **19**, 53–68.

Durre, I. and X. Yin, 2008: Enhanced radiosonde data for studies of vertical structure. *Bull. of the Amer. Met. Soc.*, **89**, 1257–1262.

Dutton, J. A., 2002: *The ceaseless wind: an introduction to the theory of atmospheric motion*. Dover Phoenix Editions, 617 pp.

Dyer, T. G. J., 1975: The assignment of rainfall stations into homogeneous groups: an application of principal component analysis. *Quarter. J. Roy. Met. Soc.*, **101**, 1005–1013.

Egger, J., 1995: Regional statistical-dynamical climate modeling: tests. *Contr. Atm. Phys.*, **85**, 281–289.

Eischeid, J., C. B. Baker, T. Karl, and H. F. Díaz, 1995: The quality control of long-term climatological data using objective data analysis. *J. Appl. Meteor.*, **34**, 2787–2795.

El-Kadi, A. K. A. and P. A. Smithson, 1992: Atmospheric classifications and synoptic climatology. *Prog. Phys. Geogr.*, **16**, 432–455.

Endlich, R. M., F. L. Ludwig, C. M. Bhumralkar, and M. A. Estoque, 1982: A diagnostic model for estimating winds at potential sites for wind turbines. *J. Appl. Meteor.*, **21**, 1441–1454.

Esteban, P., J. Martín-Vide, and M. Mases, 2006: Daily atmospheric circulation catalogue for western Europe using multivariate techniques. *Int. J. Climatol.*, **26**, 1501–1515.

Fairless, D., 2007: How did a little Spanish province become one of the world's wind-energy giants? *Nature*, **447**, 1046–1048.

Farr, T. G., P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf, 2007: The shuttle radar topography mission. *Rev. Geophys.*, **45**, RG2004, doi:10.1029/2005RG000183.

Faucher, M., W. R. Burrows, and L. Pandolfo, 1999: Empirical-statistical reconstruction of surface marine winds along the western coast of Canada. *Clim. Res.*, **11**, 173–190.

Feng, S., Q. Hu, and W. Qian, 2004: Quality control of daily meteorological data in China, 1951-2000: a new dataset. *Int. J. Climatol.*, **24**, 853–870.

Fernández, J., J. P. Montávez, J. Sáez, J. F. González-Rouco, and E. Zorita, 2007: Sensitivity of the MM5 mesoscale model to physical parameterizations for regional climate studies: annual cycle. *J. Geophys. Res.*, **112**, D04101.

Fernández-Mills, G., 1994: Principal component analysis of precitipation and rainfall regionalization in Spain. *Theor. Appl.Climatol.*, **50**, 169–183.

Fiebrich, C. A., D. L. Grimsley, R. A. McPherson, and K. A. Kesler, 2006: The value of routine site visits in managing and maintaining quality data from the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.*, **23**, 406–416.

Filippov, V. V., 1968: Quality control procedures for meteorological data. Technical report, WMO, Geneva.

Font, I., 2000: *Climatología de España y Portugal*. Ediciones Universidad de Salamanca, second edition, 422 pp.

Foufoula-Georgiou, E. and P. Kumar, eds., 1995: *Wavelets in Geophysics*. Academic Press, 373 pp.

Fovell, R. G. and M. C. Fovell, 1993: Climate Zones of the Conterminous United States Defined Using Cluster Analysis. *J. Climatol.*, **6**, 2103–2135.

Frank, H. P. and L. Landberg, 1997: Modelling the wind climate of Ireland. *Bound. Layer Meteorol.*, **85**, 359–378.

Frey-Buness, F., D. Heimann, and R. Sausen, 1995: A statistical-dynamical downscaling procedure for global climate simulations. *Theor. Appl.Climatol.*, **50**, 117–131.

Fuentes, U. and D. Heimann, 1996: Verification of statistical-dynamical downscaling in the Alpine region. *Clim. Res.*, **7**, 151–168.

— 2000: An improved statistical-dynamical downscaling scheme and its application to the Alpine precipitation climatology. *Theor. Appl.Climatol.*, **65**, 119–135.

Gadgil, S. and N. V. Joshi, 1983: Climatic clusters of the Indian region. *J. Climatol.*, **3**, 47–63.

Gandin, L. S., 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev.*, **116**, 1137–1156.

García, A., J. L. Torres, E. Prieto, and A. de Francisco, 1998: Fitting wind speed distributions: a case study. *Solar Energy*, **2**, 139–144.

García, L., 1985: La predicción del Tiempo en el Valle del Ebro. Technical Report Serie A, no. 38, INM, Madrid.

García, L. and A. Reija, 1994: *Tiempo y clima en España. Meteorología de las autonomías*. Dossat 2000, 410 pp.

García-Bustamante, E., J. F. González-Rouco, P. A. Jiménez, J. Navarro, and J. P. Montávez, 2008: The influence of the Weibull assumption in monthly wind energy estimation. *Wind Energy*, **11**, 483–502.

— 2009: A comparison of methodologies for monthly wind energy estimations. *Wind Energy*, (In Press.).

Georgelin, M. and E. Richard, 1996: Numerical simulation of flow diversion around the Pyrenees: a Tramontana case study. *Mon. Wea. Rev.*, **124**, 687–700.

Gesch, D. B. and K. S. Larson, 1996: Techniques for development of global 1-kilometer digital elevation models. *In: Pecora Thirteen, Human interactions with the Environment - Perspectives from Space*, Sioux Falls, South Dakota, August 20-22.

Giorgi, F., 1990: Simulation of regional climate using a limited area model nested in a general circulation model. *J. Climate*, **3**, 941–963.

— 2006: Regional climate modeling: status and perspectives. *J. Phys. IV France*, **139**, 101–118.

Giorgi, F. and G. T. Bates, 1989: The climatological skill of a regional model over complex terrain. *Mon. Wea. Rev.*, **117**, 2325–2347.

Giorgi, F., G. T. Bates, and S. J. Nieman, 1993: The multi-year surface climatology of a regional atmospheric model over the western United States. *J. Climate*, **6**, 75–95.

Giorgi, F. and O. Means, 1991: Aproaches to the simulation of regional climate change: a review. *Rev. Geophys.*, **29**, 191–216.

Gomez-Muñoz, V. M. and M. A. Porta-Gándara, 2002: Local wind patterns for modeling renewable energy systems by means of cluster analysis techniques. *Renew. Energy*, **2**, 171–182.

Gong, X. and M. B. Richman, 1995: On the aplication of cluster analysis to growing season precipitation data in North America East of the Rockies. *J. Climate*, **8**, 897–931.

González-Rouco, J. F., H. Heyen, E. Zorita, and F. Valero, 2000: Agreement between observed rainfall trends and climate change simulations in the southwest of Europe. *J. Climate*, **13**, 3057–3065.

González-Rouco, J. F., J. L. Jiménez, V. Quesada, and F. Valero, 2001: Quality control and homogeneity of precipitation data in teh southwest of Europe. *J. Climate*, **8**, 964–978.

Goodin, W. R., G. J. McRae, and J. H. Seinfeld, 1980: An objective analysis technique for constructing three-dimensional urban scale wind fields. *J. Appl. Meteor.*, **19**, 98–108.

Goswami, D. Y., 2008: A review and future prospects of renewable energy in the global energy system. *Advanced Technology of Electrical Engineering and Energy*, **10**, 55–62.

Grant, E. L. and R. S. Leavenworth, 1972: *Statistical Quality Control*. McGraw-Hill Book Company, New York.

Graybeal, D. Y., 2006: Relationships among daily mean and maximun wind speeds, with application to data quality assurance. *Int. J. Climatol.*, **26**, 29–43.

Graybeal, D. Y., A. T. Degaetano, and K. L. Eggleston, 2004a: Complex quality assurance of historical hourly surface airways meteorological data. *J. Atmos. Oceanic Technol.*, **21**, 1156–1169.

— 2004b: Improved quality assurance for historical hourly temperature and humidity: development and application to environmental analysis. *J. Appl. Meteor.*, **43**, 1722–1735.

Green, M., R. G. Flocchini, and L. O. Myrup, 1993: Use of Temporal Principal Components Analysis to Determine Seasonal Periods. *J. Appl. Meteor.*, **32**, 986–995.

Green, M., L. O. Myrup, and R. G. Flocchini, 1992a: A method for classification of wind field patterns and its application to southern California. *Int. J. Climatol.*, **12**, 111–135.

Green, M. C., R. G. Flocchini, and L. O. Myrup, 1992b: Relationship of the extinction coefficient distribution to wind field patterns in southern California. *Atmos. Environ.*, **26A**, 827–840.

Gregory, S., 1975: On the delimitation of regional patterns of recent climatic fluctuations. *Weather*, **30**, 276–288.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398 1 STR.

Guardans, R. and I. Palomino, 1995: Description of wind field dynamic patterns in a valley and their relation to mesoscale and synoptic-scale meteorological simulations. *J. Appl. Meteor.*, **34**, 49–67.

Gutiérrez, J. M., A. S. C. no, R. Cano, and M. A. Rodríguez, 2004: Clustering methods for statistical downscaling in short-range weather forecast. *Mon. Wea. Rev.*, **132**, 2169–2183.

Hanna, S. R. and R. Yang, 2001: Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients and mixing depths. *J. Appl. Meteor.*, **40**, 1095–1104.

Hannachi, A., I. T. Jolliffe, D. B. Stephenson, and N. Trendafilov, 2006: In search of simple structures in climate: simplifying EOFs. *Int. J. Climatol.*, **89**, 7–28.

Hardy, D., 1977: Empirical analysis of vector observations. *Geophys. Res. Lett.*, **4**, 319–320.

Hardy, D. and J. J. Walton, 1978: Principal Component Analysis of Vector Wind Mesurements. *J. Appl. Meteor.*, **17**, 1153–1162.

Heimann, D., 1986: Estimation of regional surface layer wind field characteristics using a three-layer mesoscale model. *Beitr. Phys. Atmos.*, **59**, 518–537.

— 2001: A model-based wind climatology of the eastern Adriatic coast. *Meteorol. Z.*, **10**, 5–16.

Hewitson, B. C. and R. G. Crane, 1996: Climate downscaling: techniques and application. *Clim. Res.*, **19**, 85–95.

Holton, J. R., 2004: *Convective parameterization for mesoscale models: The Kain-Fritcsh scheme, The representation of cumulus convection in numerical models*. Amer. Meteor. Soc, 535 pp.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341.

Horel, J. D., 1984: Complex principal component analysis: theory and examples. *J. Climate Appl. Meteor.*, **23**, 1660–1673.

Howard, T. and P. Clark, 2007: Correction and downscaling of NWP wind speed forecasts. *Meteor. Atmos. Phys.*, **14**, 105–116.

Hubbard, K. G., S. Goddard, W. D. Sorensen, N. Wells, and T. T. Osugi, 2005: Performance of quality assurance procedures for an applied climate information system. **22**, 105–112.

Hubbard, K. G., N. B. Guttman, J. You, and Z. Chen, 2007: An improved QC process for temperature in the daily cooperative weather observations. **24**, 206–213.

Hunt, J. C. R., S. Leibovich, and K. J. Richards, 1988: Turbulent shear flow over low hills. *Quarter. J. Roy. Met. Soc.*, **114**, 1435–1470.

Huth, R., 1996: An intercomparison of computer-assisted circulation classification methods. *Int. J. Climatol.*, **16**, 893–922.

— 2002: Statistical downscaling of daily temperature in central Europe. *J. Climate*, **15**, 1731–1742.

Jackson, P. S. and J. C. R. Hunt, 1975: Turbulent wind flow over a low hill. *Quarter. J. Roy. Met. Soc.*, **1975**, 833–851.

Janjić, Z. I., 1994: The step-mountain eta coordinate model: further developments of the convection, viscous layer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–947.

Jenkins, G. J., P. J. Mason, W. H. Moores, and R. I. Sykes, 1981: Measurements of the flow structure around Ailsa Craig, a steep, tree-dimensional, isolated hill. *Quarter. J. Roy. Met. Soc.*, **107**, 833–851.

Jenkinson, A. F. and F. P. Collison, 1977: An initial climatology of gales over North Sea. *Synoptic Climatology Branch Memorandum n^o 62*. Meteorological Office: Bracknell.

Jiang, Q., J. D. Doyle, T. Haack, M. J. Dvorak, C. L. Archer, and M. Z. Jacobson, 2008: Exploring wind energy potential off the California coast. *Geophys. Res. Lett.*, **35**, L20819.

Jiménez, P. A., J. F. González-Rouco, E. García-Bustamante, J. Navarro, J. P. Montávez, J. Vilà-Guerau de Arellano, J. Dudhia, and A. Roldan, 2009a: Surface wind regionalization over complex terrain: evaluation and analysis of a high resolution WRF numerical simulation. *J. Appl. Meteor. & Climatol.*, (Accepted).

Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, E. García-Bustamante, and J. Navarro, 2009b: Climatology of wind patterns in the northeast of the Iberian Peninsula. *Int. J. Climatol.*, **29**, 501–525.

Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, E. García-Bustamante, J. Navarro, J. Vilà-Guerau de Arellano, J. Dudhia, and A. Roldan, 2009c: High-resolution dynamical downscaling of daily mean surface circulations over complex terrain, (Manuscript in preparation).

Jiménez, P. A., J. F. González-Rouco, J. P. Montávez, J. Navarro, E. García-Bustamante, and F. Valero, 2008: Surface wind regionalization in complex terrain. *J. Appl. Meteor. & Climatol.*, **47**, 308–325.

Jiménez, P. A., J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante, 2009d: Quality control and bias correction of high reesolution surface wind observations from automated weather stations, (Manuscript in preparation).

Jiménez, P. A., J. P. Montávez, E. García-Bustamante, J. Navarro, J. M. Jiménez-Gutiérrez, E. E. Lucio-Eceiza, and J. F. González-Rouco, 2009e: Diurnal surface wind variations over complex terrain. *Física de la Tierra*, (In revision).

Johnson, S. C., 1967: Hierarchical clustering schemes. *Psychometrika*, **32**, 261–274.

Juang, H.-M. H., C.-T. Lee, Y. Zhang, Y. Song, M.-C. Wu, Y.-L. Chen, K. Komada, and S.-C. chen, 2005: Applying Horizontal Diffusion on Pressure Surface to Mesoscale Models on Terrain-Following Coordinates. *Mon. Wea. Rev.*, **133**, 1384–1402.

Jungo, P., S. Goyette, and M. Beniston, 2002: Daily wind gust speed probabilities over Switzerland according to three types of synoptic circulation. *Int. J. Climatol.*, **22**, 485–499.

Kaas, E., T.-S. Li, and T. Schmith, 1996: Statistical hindcast of wind climatology in North Atlantic and northwestern European region. *Clim. Res.*, **110**, 97–110.

Kaihatu, J. M., R. A. Handler, G. O. Marmorino, and L. K. Shay, 1998: Empirical orthogonal function analysis of ocean surface currents using complex and real-vector methods. *J. Atmos. Oceanic Technol.*, **15**, 927–941.

Kain, J. S. and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802.

— 1993: *Convective parameterization for mesoscale models: The Kain-Fritcsh scheme, The representation of cumulus convection in numerical models*. Amer. Meteor. Soc, 246 pp.

Kaiser, H. F., 1958: The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187–200.

Källstrand, B., H. Bergström, Jøjstrup, and A.-S. Smedman, 2000: Mesoscale wind field modifications over the Baltic Sea. *Bound. Layer Meteorol.*, **95**, 161–188.

Kalnay et al., 1996: The NCEP/NCAR 40 year reanalysis project. *Bull. of the Amer. Met. Soc.*, **77**, 437–471.

Kanamitsu, M. and H. Kanamaru, 2007: Fifty-seven-year California reanalysis downscaling at 10 km CaRD10. part I: system detail and validation with observations. *J. Climate*, **20**, 5533–5571.

Karl, T. R., J. D. Tarpley, R. G. Quayle, H. F. Diaz, D. A. Robinson, and R. S. Bradley, 1989: The recent climate record: what it can and cannot tell us. *Rev. Geophys.*, **27**, 405–430.

Karl, T. R. and C. N. Williams, 1987: An approach to adjusting climatological time series for discontinuous inhomogeneities. *Climate Appl. Meteor.*, **26**, 1744–1763.

Karl, T. R., C. N. Williams, P. J. Young, and W. M. Wendland, 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum, and mean temperature for the United States. *Climate Appl. Meteor.*, **26**, 145–160.

Kastendeuch, P. P. and P. Kaufmann, 1997: Classification of summer wind fields over complex terrain. *Int. J. Climatol.*, **17**, 521–534.

Kaufmann, P. and R. O. Weber, 1996: Classification of Mesoscale Wind Fields in the MISTRAL Field Experiment. *J. Appl. Meteor.*, **35**, 1963–1979.

— 1998: Directional correlation coefficient for channeled flow and application to wind data over complex terrain. *J. Atmos. Oceanic Technol.*, **15**, 89–97.

Kaufmann, P. and C. D. Whiteman, 1999: Cluster-Analysis Classification of Wintertime Wind Patterns in the Grand Canyon Region. *J. Appl. Meteor.*, **38**, 1131–1147.

Kidson, J. W., 1994: An automated procedure for the identification of synoptic types applied to the New Zealand region. *Int. J. Climatol.*, **14**, 711–721.

Kim, J. W., T. K. Chang, N. L. Bajer, D. S. Wilks, and W. L. Gates, 1984: The statistical problem of climate inversion: determination of the relationship between local and large-scale climate. *Mon. Wea. Rev.*, **11**, 2069–2977.

Kirchhofer, W., 1974: Classification of European 500 mb patterns. Technical Report 43, Swiss Meteorological Institute, Zurich.

Klemp, J. B. and R. Wihelmson, 1978: The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.*, **35**, 1070–1096.

Klink, K., 2002: Trends and interannual variability of wind speed distributions in Minesota. *J. Climate*, **15**, 3311–3317.

Klink, K. and C. J. Willmott, 1989: Principal components of the surface wind field in the United States: a comparison of analyses based upon wind velocity, direction, and speed. *Int. J. Climatol.*, **9**, 293–308.

Krishnamurti, T. N. and L. Bounoua, 1995: *An introduction to numerical weather prediction techniques*. TF-CRC, 293 pp.

Kunkel, K. E., K. Andsager, G. Conner, W. L. Decker, H. J. Hillaker, Jr., P. Naber Knox, F. V. Nurnberger, J. C. Rogers, K. Scheeringa, W. M. Wendland, J. Zandlo, and J. R. Angel, 1998: An expanded digital daily database for climatic resources applications in the midwestern United States. *Bull. of the Amer. Met. Soc.*, **79**, 1357–1366.

Kunkel, K. E., D. R. Easterling, K. Hubbard, K. Redmond, K. Andsager, M. C. Kruk, and M. L. Spinar, 2005: Quality control of pre-1948 cooperative observer netweork data. *J. Atmos. Oceanic Technol.*, **22**, 1691–1704.

Lamb, H. H., 1972: British Isles weather types and a register of the daily sequence of circulation patterns, 1861-1971. *Geophysical Memoirs* **116**, London.

Laprise, R., 1992: The Euler equations of motion with hydrostatic pressure as an independent variable. *Mon. Wea. Rev.*, **120**, 197–207.

Lazante, J. R., 1996: Resistant, robust, and nonparametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *Int. J. Climatol.*, **22**, 1197–1226.

Leduc, M. and R. Laprise, 2009: Regional climate model sensitivity to domain size. *Clim. Dyn.*, **32**, 833–854.

Lewis, J. M., 1998: Clarifying the dynamics of the general circulation: Phillips's 1956 experiment. *Bull. of the Amer. Met. Soc.*, **79**, 39–60.

Li, Y. and I. Smith, 2009: A statistical downscaling model for southern Australia winter rainfall. *J. Climate*, **22**, 1142–1158.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092.

Linderson, M., 2001: Objective classification of atmospheric circulation over southern Scandinavia. *Int. J. Climatol.*, **21**, 155–169.

Linés, A., 1981: Perturbaciones típicas que afectan a la península Ibérica y precipitaiones asociadas. Madrid: Instituto Nacional de Meteorología.

Lo, J. C.-F., Z.-L. Yang, and R. A. Pielke Sr., 2008: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model. *J. Geophys. Res.*, **113**, D09112.

Lorentz, E. N., 1967: The nature and theory of the general circulation of the atmosphere. Technical report, WMO, Geneva.

Lu, X., M. B. McElroy, and J. Kiviluoma, 2009: Global potential for wind-generated electricity. *Proc. Natl. Acad. Sci. USA*, **Published online 2009 June 19**, doi: 10.1073/pnas.0904101106.

Ludwig, F. L., J. Horel, and C. D. Whiteman, 2004: Using EOF analysis to identify important surface winds patterns in mountain valleys. *J. Appl. Meteor.*, **43**, 969–983.

Ludwig, F. L., D. K. Miller, and S. G. Gallaher, 2006: Evaluationg a hibrid pognostic-diagnostic model that improves wind forecast resolution in complex coastal topography. *J. Appl. Meteor. & Climatol.*, **45**, 155–177.

Lund, I. A., 1963: Map-Pattern Classification by Statistical Methods. *J. Appl. Meteor.*, **2**, 56–65.

Mahrer, Y. and R. A. Pielke, 1977: The effects of topography on sea and land breezes in a two-dimensional numerical model. *Mon. Wea. Rev.*, **105**, 1151–1162.

Mahrer, Y., M. Segal, and R. Pielke, 1985: Mesoscale modelling of wind energy over non-homogeneous terrain. *Bound. Layer Meteorol.*, **31**, 13–23.

Mallants, D. and J. Feyen, 1990: Defining homogeneous precipitaion regions by means of principal component analysis. *J. Appl. Meteor.*, **29**, 892–901.

Malmgren, B. A. and A. Winter, 1999: Climate zonation in Puerto Rico based on principal components analysis and an artificial neural network. *J. Climate*, **12**, 977–985.

Manabe, S. and T. B. Terpstra, 1974: The effects of mountains on the general circulation of the atmosphere as identified by numerical experiments. *J. Atmos. Sci.*, **31**, 3–42.

Mardia, R. V. and P. E. Jupp, 1999: *Directional statistics*. J. Wiley and Sons, 429 pp.

Marinucci, M. R., F. Giorgi, M. Beniston, M. Wild, P. Tschuck, A. Ohmura, and A. Bernasconi, 1995: High resolution simulations of January and July climate over the western alpine region with a nested regional modeling system. *Theor. Appl.Climatol.*, **51**, 119–138.

Martín-Vide, J., 2001: Limitations of an objective weather-typing system for the iberian peninsula. *Weather*, **56**, 248–250.

Martner, B. E. and J. D. Marwitz, 1982: Wind characteristics in Southern Wyoming. *J. Appl. Meteor.*, **21**, 1815–1827.

Mason, P. J. and J. C. King, 1984: Atmospheric flow over a succession of nearly two-dimensional ridges and valleys. *Quarter. J. Roy. Met. Soc.*, **110**, 821–845.

— 1985: Measurements and predictions of flow and turbulence over an isolated hill of moderate slope. *Quarter. J. Roy. Met. Soc.*, **111**, 617–640.

Mason, P. J. and R. I. Sykes, 1979: Flow over an isolated hill of moderate slope. *Quarter. J. Roy. Met. Soc.*, **105**, 383–395.

Mass, C. F. and Y.-H. Kuo, 1998: Regional real-time numerical weather prediction: current status and future potential. *Bull. of the Amer. Met. Soc.*, **79**, 253–263.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. of the Amer. Met. Soc.*, **83**, 407–430.

Masson, V. and P. Bougeault, 1996: Numerical simulation of a low-level wind created by complex orography: a Cierzo case study. *Mon. Wea. Rev.*, **124**, 701–715.

McGregor, J. L., 1997: Regional climate modeling. *Meteor. Atmos. Phys.*, **63**, 1599–1628.

McPherson, R. A., C. A. Fiebrich, K. C. Crawford, R. L. Elliott, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A. K. ad S. J. Stadler, A. D. Melvin, A. J. Sutherland, H. Shrivastava, J. D. Carlson, J. M. Wolfinbarger, J. P. Bostic, and D. B. Demko, 2007: State wide monitoring of the mesoscale environment: a technical update on the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.*, **24**, 301–321.

McVicar, T. R., T. G. Van Niel, L. T. Li, M. L. Roderick, D. P. Rayner, L. Ricciardulli, and R. J. Donohue, 2008: Wind speed climatology and trends for Australia, 1975-2006: capturing the stilling phenomenon and comparison with near-suface reanalysis output. *Geophys. Res. Lett.*, **35**, L20403, doi:10.1029/2008GL035627.

Meek, D. W. and J. L. Hatfield, 1994: Data quality checking for single station meteorological databases. *Agricultural and Forest Meteorology*, **69**, 85–109.

Mellor, G. L. and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875.

Mengelkamp, H.-T., 1999: Wind climate simulation over complex terrain and wind turbine energy output estimation. *Theor. Appl.Climatol.*, **63**, 129–139.

Mengelkamp, H.-T., H. Kapitza, and U. Pflüger, 1997: Statistical-dynamical downscaling of wind climatologies. *J. Wind Eng. Ind. Aerodyn.*, **67&68**, 449–457.

Miao, J.-F., D. Chen, K. Wyser, K. Borne, J. Lindgren, M. K. S. Strandevall, S. Thorsson, C. Achberger, and E. Almkvist, 2008: Evaluation of MM5 mesoscale model at local scale for air quality applications over the Swedish west coast: Influence of PBL and LSM parameterizations. *Meteor. Atmos. Phys.*, **99**, 77–103.

Michelson, S. A. and J.-W. Bao, 2008: Sensitivity of low-level winds simulated by the WRF model in California's Central Valley to uncertainties in the large-scale forcing and soil initialization. *J. Appl. Meteor. & Climatol.*, **47**, 3131–3149.

Millán, M. M., M. J. Sanz, R. Salvador, and E. Mantilla, 2003: Atmospheric dynamics and ozone cycles related to nigrogen deposition in the western mediterranean. *Global Planetary Change*, **118**, 167–186.

Milligan, G. W., 1980: An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, **50**, 325–342.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682.

Muñoz-Díaz, D. and F. S. Rodrigo, 2004: Spatio-temporal patterns of seasonal rainfall in Spain (1912-200) using cluster and principal component analysis: comparison. *Ann. Geophys.*, **22**, 1435–1448.

Najac, J., J. Boéa, and L. Terray, 2009: A multi-model ensemble approach for assessment of climate change impact on surface winds in France. *Clim. Dyn.*, **32**, 615–634.

Noguer, M., 1994: Using statistical techniques to deduce local climate distributions. an application for model validation. *Meteorol. Appl.*, **1**, 277–287.

Ogallo, L. J., 1989: The spatial and temporal patterns of the east-african seasonal rainfall derived from principal component analysis. *Int. J. Climatol.*, **9**, 145–167.

Ooyama, K. V., 1990: A thermodynamic foundation for modeling the moist atmosphere. *J. Atmos. Sci.*, **47**, 2580–2593.

Oreskes, N., 1998: Evaluation (not validation) of quantitative models. *Environ Health Perspectives*, **106**, 1453–1459.

Oreskes, N., K. Shrader-Frechette, and K. Belitz, 1994: Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, **263**, 641–646.

Palmén, E. and L. A. Vuorela, 1963: On the mean meridional circulations in the northern hemisphere during the Winter season. *Quarter. J. Roy. Met. Soc.*, **89**, 131–138.

Palutikof, J. P., P. M. Kelly, T. D. Davies, and J. A. Halliday, 1987: Impacts of Spatial and Temporal Windspeed Variability on Wind Energy Output. *J. Appl. Meteor.*, **26**, 1124–1133.

Pan, Z., E. Takle, W. Gutowski, and R. Turner, 1999: Long simulation of regional climate as a sequence of short segments. *Mon. Wea. Rev.*, **127**, 308–321.

Pandžić, K. and T. Likso, 2005: Eastern Adriatic typical wind field patterns and large-scale atmospheric conditions. *Int. J. Climatol.*, **25**, 81–98.

Parker, D. E., T. A. Basnett, S. J. Brown, M. Gordon, E. B. Horton, and N. A. Rayner, 2000: Climate observations - the instrumental record. *Space Sci. Rev.*, **94**, 309–320.

Paul, S., C. M. Liu, J. M. Chen, and S. H. Lin, 2008: Development of a statistical downscaling model for projecting monthly rainfall over east asia from a general circulation model output. *J. Geophys. Res.*, **113**, D15117.

Paulsen, B. M. and J. L. Schoeder, 2005: An examination of tropical and extratropical gust factors and the associated wind speed histograms. *J. Appl. Meteor.*, **44**, 1197–1226.

Pérez-Landa, G., P. Ciais, M. J. Ciais, M. J. Sanz, B. Gioli, F. Miglietta, J. L. Palau, G. Gangoiti, and M. M. Millan, 2007: Mesoscale circulations over complex terrain in the Valencia coastal region, Spain - part 1: simulation of diurnal circulation regimes. *Atmos. Chem. Phys. Discuss.*, **7**, 1825–1849.

Peterson, T. C., D. R. Easterling, T. R. Karl, P. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, G. D, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szentimrey, J. Salinger, E. J. Førland, I. Hanssen-Bauer, H. Alexandersson, P. Jones, and D. Parker, 1998a: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517.

Peterson, T. C., R. Vose, R. Schmoyer, and V. Razuvaev, 1998b: Global Historical Climatology Network (GHCN) quality control of monthly temperature data. *Int. J. Climatol.*, **18**, 1169–1179.

Phillips, N. A., 1956: The general circulation of the atmosphere: a numerical experiment. *Quarter. J. Roy. Met. Soc.*, **82**, 123–164.

Physick, W. L., 1988: Review: mesoscale modelling in complex terrain. *Earth-Sci. Rev*, **25**, 199–235.

Pielke, R. A., 1985: The use of mesoscale numerical models to assess wind distribution and boundary-layer structure in complex terrain. *Bound. Layer Meteorol.*, **31**, 217–231.

— 2002: *Mesoscale meteorological modeling*. Academic Press, 676 pp.

Pielke Sr., R., C. A. Davey, D. Niyogi, S. Fall, J. Steinweg-Woods, K. Hubbard, X. Lin, M. Cai, Y.-K. Lim, H. Li, J. Nielsen-Bammon, K. Gallo, R. Hale, R. Mahmood, S. Foster, R. T. McNider, and P. Blamken, 2007a: Unresolved issues with the assessment of multidecadal global land surface temperature trends. *J. Geophys. Res.*, **112**, D24S08, doi:10.1029/2006JD008229.

Pielke Sr., R., J. Nielsen-Gammon, C. Davey, J. Angel, O. Bliss, N. Doesken, M. Cai, S. Fall, D. Niyogi, K. Gallo, R. Hale, K. G. Hubbard, X. Lin, H. Li, and S. Raman,

2007b: Documentation of uncertainties and biases associated with surface temperature measurement sites for climate change assessment. *Bull. of the Amer. Met. Soc.*, **88**, 913–928.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, 425 pp.

Prieto, L., R. García-Herrera, J. Díaz, E. Hernández, and T. del Teso, 2004: Minimum extreme temperatures over Peninsular Spain. *Global Planetary Change*, **4**, 59–71.

Pryor, S. C., R. J. Barthelmie, and E. Kjellström, 2005a: Potential climate change impact on wind energy resources in norhtern Europe: analyses using a regional climate model. *Clim. Dyn.*, **25**, 815–835.

Pryor, S. C., J. T. Schoof, and R. J. Barthelmie, 2005b: Climate change impacts on wind speeds and wind energy density in northern Europe: empirical downscaling of multiple AOGCMs. *Clim. Res.*, **29**, 183–198.

— 2005c: Empirical downscaling of wind speed probability distributions. *J. Geophys. Res.*, **110**.

— 2006: Winds of change?: Proyections of near-surface winds under climate change scenarios. *Geophys. Res. Lett.*, **33**, L11702.

Qian, J.-H., A. Seth, and S. Zebiak, 2003: Reinitialized versus continuous simulations for regional climate downscaling. *Mon. Wea. Rev.*, **131**, 2857–2874.

Queney, P., 1948: The problem of air flow over mountains: a summary of theoretical studies. *Bull. of the Amer. Met. Soc.*, **29**, 16–26.

Ratto, C. F., R. Festa, C. Romeo, O. A. Frumento, and M. Galluzzi, 1994: Mass-consistent models for wind fields over complex terrain - the state- of-the-art. *Env. Soft.*, **9**, 247–268.

Rauscher, S. A., A. Seth, J.-H. Qian, and S. J. Camargo, 2006: Domain choice in an experimental nested modeling prediction system for South America. *Theor. Appl.Climatol.*, **86**, 229–246.

Reek, T., S. R. Doty, and T. W. Owen, 1992: A deterministic approach to validate of historical daily temperature and precipitation data from the cooperative network. *Bull. of the Amer. Met. Soc.*, **30**, 511–523.

Reid, S. and R. Turner, 2001: Correlation of real and model wind speeds in different terrains. *Wea. Forecasting*, **16**, 620–627.

Reynolds, O., 1895: On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Phil. Trans. R. Soc. Lond.*, **186**, 123–164.

Richardson, F. L., 1922: *Weather prediction by numerical process*. Cambridge University Press.

Richman, M. B., 1986: Rotation of principal components. *J. Climatol.*, **6**, 293–335.

Rife, D. R. and C. A. Davis, 2005: Verification of temporal variations in mesoscale numerical wind forecast. *Mon. Wea. Rev.*, **133**, 3368–3381.

Rife, D. R., C. A. Davis, Y. Liu, and T. T. Warner, 2004: Predictability of Low-Level Winds by Mesoscale Meteorological Models. *Mon. Wea. Rev.*, **132**, 2533–2569.

Rodríguez, S., X. Querol, A. Alastuey, G. Kallos, and O. Kakaliagou, 2001: Saharan dust contributions to PM10 and TSP levels in Southern and Eastern Spain. *Atmos. Environ.*, **35**, 2433–2447.

Romero, R., C. Ramis, and J. A. Guijarro, 1999a: Daily rainfall patterns in the Spanish Mediterranean area: an objective classification. *Int. J. Climatol.*, **19**, 95–112.

Romero, R., C. Ramis, J. A. Guijarro, and G. Summer, 1999b: Daily rainfall affinity areas in Mediterranean Spain. *Int. J. Climatol.*, **19**, 557–578.

Romero, R., G. Summer, C. Ramis, and A. Genovés, 1999c: A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *Int. J. Climatol.*, **19**, 765–785.

Rontu, L., 2006: A study on parameterization of orography-related momentum fluxes in a synoptic-scale NWP model. *Tellus*, **58A**, 69–81.

Sahsamanoglou, H. S., 1990: A contribution to the study of action centres in the North Atlantic. *Int. J. Climatol.*, **10**, 247–261.

Sailor, D. J., T. Hu, X. Li, and J. N. Rosen, 2000: A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change. *Renew. Energy*, **19**, 359–378.

Sanchez-Lorenzo, A., M. Brunetti, J. Calbo, and J. Martín-Vide, 2007: Recent spatial and temporal variability and trends of sunshine duration over the Iberian Peninsula from a homogenized data set. *J. Geophys. Res.*, **112**, D20115–1–18.

Sanchez-Lorenzo, A., J. Calbo, and J. Martín-Vide, 2008: Spatial and temporal trends in shunshine duration over western Europe (1938-200). *J. Climate*, **2008**, 6089–6098.

Sandström, S., 1997: Simulations of the climatological wind field in the Baltic sea area using a mesoscale higher-order closure model. *J. Appl. Meteor.*, **36**, 1541–1552.

Schneider, T., 2006: The general circulation of the atmosphere. *Annu. Rev. Earth Planet. Sci.*, **34**, 655–688.

Scorer, R. S., 1949: Theory of waves in the lee of mountains. *Quarter. J. Roy. Met. Soc.*, **76**, 41–56.

Seneviratne, S. I., R. D. Koster, Z. Guo, P. A. Dirmeyer, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, D. Mocko, K. W. Oleson, and D. Verseghy, 2006: Soil moisture memory in AGCM simulations: analysis of global land-atmosphere coupling experiemnt (GLACE) data. *J. hydrometeor.*, **7**, 1090–1112.

Seth, A. and F. Giorgi, 1998: The effects of domain choice on summer precipitation simulation and sensitivity in a regional climate model. *J. Climate*, **11**, 2698–2712.

Shafer, M. A., C. A. Fiebrich, D. S. Arndt, S. E. Fredrickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma mesonetwork. *J. Atmos. Oceanic Technol.*, **17**, 474–494.

Sherman, C. A., 1978: A mass-consistent model for wind fields over complex terrain. *J. Appl. Meteor.*, **17**, 312–319.

Simmons, A. J. and J. K. Gibson, 2000: The ERA-40 project plan, ERA-40 Project report series no. 1. Technical report, Eur. Cent. for Medium-Range Weather Forecast, Reading, UK.

Simpson, J. E., 1994: *Sea breezes and local winds*. Cambridge university press.

Skamarock, W. C. and J. B. Klemp, 1992: The stability of time-split numerical methods for hydrostatic and the nonhydrostatic elastic equations. *Mon. Wea. Rev.*, **120**, 2109–2127.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the advanced research WRF Version 2. Technical Report TN-468+STR, NCAR.

Soler, A. M., 1977: *Situaciones meteorológicas locales típicas: su persistencia y parámetros o variables más característicos*. Ph.D. thesis, UCM.

Soriano, C., A. Fernández, and J. Martin-Vide, 2006: Objective synoptic classification combined with high resolution meteorological models for wind mesoscale studies. *Meteor. Atmos. Phys.*, **91**, 165–181.

Spellman, G., 2000: The application of an objective weather-typing system to the Iberian peninsula. *Weather*, **55**, 375–385.

Steinacker, R., M. Ratheiser, B. Bica, B. Chimani, M. Dorninger, W. gepp, C. Lotteraner, S. Schneider, and S. Tschanmet, 2006: A mesoscale data analysis and downscaling method over complex terrain. *Mon. Wea. Rev.*, **134**, 2758–2771.

Stensrud, D. J., 2007: *Parameterization schemes: keys to understanding numerical weather prediction models*. Cambridge University Press, 478 pp.

Stooksbury, D. E. and P. J. Michaels, 1991: Cluster analysis of sourtern U.S. climate stations. *Theor. Appl.Climatol.*, **44**, 143–150.

Stull, R. B., 1988: *An introuduction to boundary layer meteorology*. Kluwier academic publishers, 666 pp.

Sykes, R. I., 1980: An asymptotic theory of incompressible turbulent boundary-layer over a small hump. *J. Fluid Mech.*, **101**, 647–670.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106D7**, 7183–7192.

Taylor, P. A. and H. W. Teunissen, 1987: The Askervin hill project: overview and background data. *Bound. Layer Meteorol.*, **39**, 15–39.

Thompson, D., J. J. Kennedy, J. M. Wallace, and P. D. Jones, 2008: A large discontinuity in the mid-twentieth century observed global-mean surface temperature. *Nature*, **453**, 646–649.

Thuillier, R. H., 1987: Real-time analysis of local wind patterns for application to nuclear-emergency response. *Bull. of the Amer. Met. Soc.*, **68**.

Torrence, C. and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. of the Amer. Met. Soc.*, **79**, 61–78.

Torres, J. L., A. García, M. D. Blas, and A. D. Francisco, 2005: Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy*, **79**, 65–77.

Torres, J. L., A. García, E. Prieto, and A. De Francisco, 1999: Characterization of wind speed data according to wind direction. *Solar Energy*, **66**, 57–64.

Torres, R., E. D. Barton, P. Miller, and E. Fanjul, 2003: Spatial patterns of wind and sea surface temperature in the Galician upwelling region,. *J. Geophys. Res.*, **108(C4)**, 3130.

Trenberth, K. E. and D. A. Paolino, 1980: The northern hemispheric sea-level pressure data set: trends, errors and discontinuities. *Mon. Wea. Rev.*, **108**, 855–872.

Trigo, R. M. and C. C. DaCamara, 2000: Circulation weather types and their influence on the precipitation regime in Portugal. *Int. J. Climatol.*, **20**, 1559–1581.

Troen, I. and E. L. Petersen, 1989: *The European Wind Atlas*. Risoe National Lab., Roskilde, Denmark.

Tucker, D. F., 1997: Surface mesonets of the western United States. *Bull. of the Amer. Met. Soc.*, **78**, 1485–1495.

Uppala, S. M., P. W. Kallberg, A. J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. Janssen, R. Jenne, A. P. McNally, J. Mahfouf, J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen, 2005: The era-40 re-analysis. *Quarter. J. Roy. Met. Soc.*, **131**, 2961–3012.

Vannttsem, S. and F. Chomé, 2005: One-way nested regional climate simulations and domain size. *J. Climate*, **18**, 229–233.

Vejen, F., C. Jacobsson, U. Fredriksson, M. Moe, L. Andresen, E. Hellsten, P. Rissanen, P. Pálsdóttir, and P. Arason, 2002: Quality Control of Meteorological Observations. Automatic Methods Used in the Nordic Countries. Technical Report 8/2002, Nordklim, Nordic co-operation whithin climate activities, KLIMA.

Venables, W. N. and B. D. Ripley, 1999: *Modern applied statistics with S-plus*. Springer: New York, 3 edition.

Verburg, P. and R. E. Hecky, 2003: Wind patterns, evaporation, and related physical variables in Lake Tanganyika, east Africa. *J. Great Lakes Res.*, **29**, 48–61.

Verdin, K. L. and S. K. Greenlee, 1996: Development of continental scale digital elevation models and extraction of hydrographic features. *In: Proceedings, Third International Coonference/Workshop on integrating GIS and Environmental Modeling*, Santa Fe, Nuevo Mexico, January 21-26. National Center for Geographic Information and Analysis, Santa Barbara, California.

Vincent, L. A., X. Zhang, B. R. Bonsal, and W. D. Hogg, 2002: Homogenization of daily temperatures over Canada. *J. Climate*, **24**, 1322–1334.

Volmer, J. P., M. Deque, and D. Rousselet, 1984: EOF analysis of 500 mb geopotential: a comparison between simulation and reality. *Tellus*, **36A**, 336–347.

von Storch, H., 1995: Inconsistencies at the interface of climate impact studies and global climate research. *Meteorol. Z.*, **4 NF**, 72–80.

von Storch, H., E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *J. Climate*, **6**, 1161–1171.

von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 499 pp.

Vosper, S. B., S. D. Mobbs, and B. A. Gardiner, 2002: Measurements of the near-surface flow over a hill. *Quarter. J. Roy. Met. Soc.*, **128**, 2257–2280.

Vrac, M., M. Stein, and K. Hayhoe, 2007: Statistical downscaling of precipitation through nonhomogeneous stochastic weather typing. *Clim. Res.*, **34**, 169–184.

Vuorela, L. A. and I. Tuominen, 1964: On the mean zonal and meridional circulations and the flux of moisture in the northern hemisphere during the Summer season. *Pure and Appl. Geophys.*, **57**, 167–180.

Wade, C. G., 1987: A quality control program for surface meteorological data. *J. Atmos. Oceanic Technol.*, **4**, 435–453.

Wagner, A., 1938: Theorie und Beobachtung der periodischen Gebirgswinde. *Gerl. Beitr. Geophys.*, **52**, 408–449.

Walmsley, J. L., J. R. Salmon, and P. A. Taylor, 1982: On the application of a model of boundary-layer flow over low hills to real terrain. *Bound. Layer Meteorol.*, **23**, 17–46.

Walsh, K. and J. L. McGregor, 1995: January and July climate simulations over the Australian region using a limited area model. *J. Climate*, **8**, 2387–2403.

Walter, A., K. Keuler, D. Jacob, R. Knoche, A. Block, S. Kotlarski, G. Müller-Westermeier, D. Rechid, and W. Ahrens, 2006: A high resolution data set of German wind velocity 1951-2001 and comparison with regional climate model results. *Meteorol. Z.*, **15**, 585–596.

Wan, H., X. X. L. Wang, and V. R. Swail, 2007: A quality assurance system for canadian hourly pressure data. *J. Appl. Meteor. & Climatol.*, **46**, 1904–1817.

Wang, Y., L. R. Leung, J. L. McGregor, D.-K. Lee, W.-C. Wang, Y. Ding, and F. Kimura, 2004: Regional climate modeling: progress, challenges, and prospects. *J. Meteorol. Soc. Japan*, **82**, 1599–1628.

Ward, J. H., 1963: Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.*, **58**, 236–244.

Warner, T. T., 1989: Mesoscale atmospheric modeling. *Earth-Sci. Rev*, **26**, 221–251.

Warner, T. T., R. A. Peterson, and R. E. Treadon, 1997: A tutorial on the lateral boundary conditions as a basic potentially serious limitation to regional numerical weather prediction. *Bull. of the Amer. Met. Soc.*, **78**, 2599–2617.

Weber, R. O., 1998: Climatology of regional flow patterns around Basel. *Theor. Appl.Climatol.*, **59**, 13–27.

Weber, R. O. and M. Furger, 2001: Climatology of near-surface wind patterns over Switzerland. *Int. J. Climatol.*, **21**, 809–827.

Weber, R. O. and P. Kaufmann, 1995: Automated classification scheme for wind fields. *J. Appl. Meteor.*, **78**, 1133–1141.

Weggel, J. R., 1999: Maximum daily wind gusts related to mean daily wind speed. *Journal of Structural Engineering*, **125**, 465–468.

Wendland, W. M., 1982: Wind power as an electrical power source in Illinois. *J. Appl. Meteor.*, **21**, 423–428.

White, D., M. Richman, and B. Yarnal, 1991: Climate regionalization and rotation of principal components. *Int. J. Climatol.*, **11**, 1–25.

Whiteman, C. D., 2000: *Mountain meteorology: fundamentals and applications*. Oxford University Press, 355 pp.

Wicker, L. J. and W. C. Skamarock, 2002: Time splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, **130**, 2088–2097.

Wigley, T., P. D. Jones, K. R. Briffa, and G. Smith, 1990: Obtaining subgrid scale information from coarse-resolution general circulation model output. *J. Geophys. Res.*, **95**, 1943–1953.

Wijngaard, J. B., A. M. Klein Tank, and G. P. Können, 2003: Homogeneity of 20th century european daily temperature and precipitation series. *Int. J. Climatol.*, **23**, 679–692.

Wilby, R. L. and T. Wigley, 1997: Downscaling general circulation model output: a review of methods and limitations. *Prog. Phys. Geog.*, **21**, 530–548.

Wilby, R. L., T. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks, 1998: Statistical downscaling of general circulation model output: A comparison of methods. *Water Resour. Res.*, **34**, 2995–3008.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Willmott, C. J., 1987: Synoptic weather-map classification: correlation versus sum-of-square. *Professional Geographer*, **39**, 205–297.

Wippermann, F. and G. Gross, 1981: On the construction of orographically influenced wind roses for given distributions of the large-scale wind. *Beitr. Phys. Atmos.*, **54**, 492–501.

WMO, 2008: Guide to meteorological instruments and methods of observation. Technical report, WMO-No. 8, Geneva.

Wu, H., K. G. Hubbard, and J. You, 2005: Some concerns when using data from the cooperative weather station networks: A Nebraska case study. *J. Atmos. Oceanic Technol.*, **22**, 592–602.

Xoplaki, E., J. F. González-Rouco, D. Gyalistras, J. Luterbacher, R. Rickli, and H. Wanner, 2003a: E. xoplaki and j. f. gonzález-rouco and d. gyalistras and j. luterbacher and r. rickli and h. wanner. *Clim. Dyn.*, **20**, 537–554.

Xoplaki, E., J. F. González-Rouco, and J. Luterbacher, 2003b: Mediterranean summer air temperature variability and its connection to the large-scale atmospheric circulation and SSTs. *Clim. Dyn.*, **20**, 723–739.

Xoplaki, E., J. F. González-Rouco, J. Luterbacher, and H. Wanner, 2004: Wet season Mediterranean precipitation variability: influence of large-scale dynamics and predictability. *Clim. Dyn.*, **23**, 63–78.

Yang, Y., Y.-L. Chen, and F. M. Fujioka, 2008: Effects of Trade-wind strength and direction on the leeside circulations and rainfall of the island of Hawaii. *Mon. Wea. Rev.*, **136**, 4799–4188.

Yarnal, B., 1993: *Synoptic climatology in environmental analysis*. Belhaven Press, 195 pp.

Yarnal, B., A. C. Comrie, B. Frakes, and D. P. Brown, 2001: Developments and prospects in synoptic climatology. *Int. J. Climatol.*, **21**, 1923–1950.

You, J., K. G. Hubbard, S. Nadarajah, and K. Kunkel, 2007: Performance of quality assurance procedures on daily precipitation. *J. Atmos. Oceanic Technol.*, **24**, 821–834.

Zagar, N., M. Zagar, J. Cedilnik, G. Gregoric, and J. Rakovec, 2006: Validation of mesoscale low-level winds obtained by dynamical downscaling of ERA40 over complex terrain. *Tellus*, **58A**, 445–455.

Zamora, R. J., S. Solomon, E. G. Dutton, J. W. Bao, M. Trainer, R. W. Portmann, A. B. White, D. W. Nelson, and R. T. McNider, 2003: Comparing MM5 radiative fluxes with observations gathered during the 1995 and 1999 Nashville southern oxidants studies. *J. Geophys. Res.*, **108(D2)**, 4050.

Zängl, G., 2002: An improved method for computing horizontal diffusion in a sigma-coordinate model and its application to simulations over mountainous topography. *Mon. Wea. Rev.*, **130**, 1423–1432.

Zhang, D.-L. and W.-Z. Zheng, 2004: Diurnal cycles of surface winds and temperatures as simulated by five boundary layer parametrizations. *J. Appl. Meteor.*, **43**, 157–169.

Zhang, X., X. L. Wang, and J. Corte-Real, 1997: On the relationships between daily circulation patterns and precipitation in portugal. *J. Geophys. Res.*, **102**, 13,495–13,507.

Zhong, S. and J. Fast, 2003: An Evaluation of the MM5, RAMS, and Meso-Eta models at subkilometer resolution using VTMX field campaign in the Salt Lake valley. *Mon. Wea. Rev.*, **131**, 1301–1322.

Zimmerschied, W., 1949: Acerca de las situaciones típicas de tiempo de la Península Ibérica. Technical Report Serie A, no. 20, INM, Madrid.

Zorita, E., V. Kharin, and H. von Storch, 1992: The atmospheric circulation and sea surface temperature in the North Atlantic area in winter: their interaction and relevance for Iberian precipitation. *J. Climate*, **5**, 1097–1108.

Zurański, J. A. and B. Jaśpińka, 1996: Directional analysis of extreme wind speeds in Poland. *J. Wind Eng. Ind. Aerodyn.*, **65**, 13–20.

9

Appendix A

The relationship between the typical large scale situations over the IP, the PPs, and the typical surface flows over the CFN, the WPs, was analyzed in Chapter 6. Clear associations between the PPs and WPs were found, see Table 6.2. This appendix displays the SLP anomalies for each one of the main associations (in bold in Table 6.2).

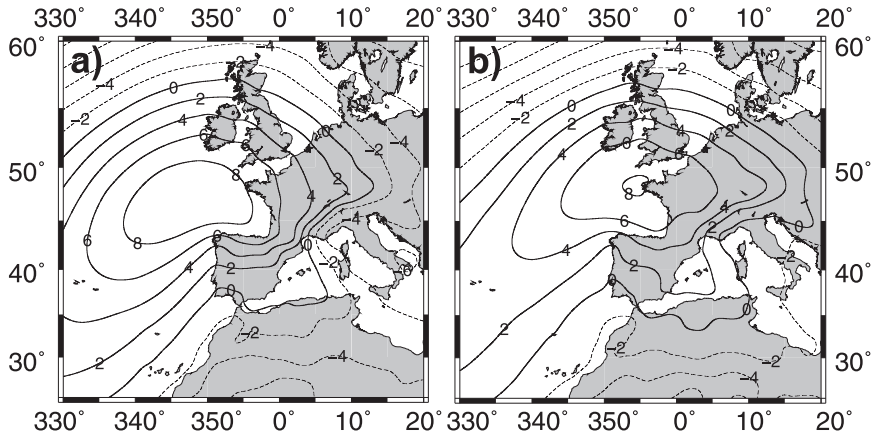

Fig. 9.1: Mean anomaly SLP fields (contour lines)for days from PP1 classified as WP1 (a) and WP2 (b).

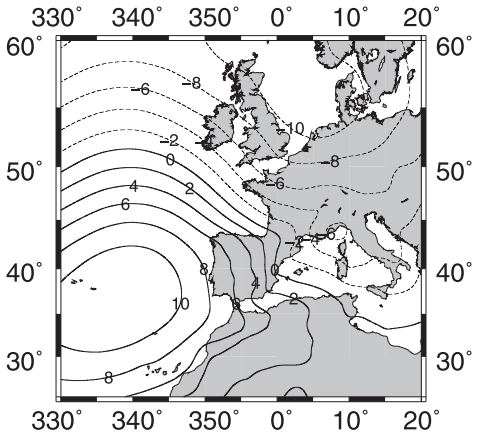

Fig. 9.2: Mean anomaly SLP fields (contour lines)for days from PP2 classified as WP3.

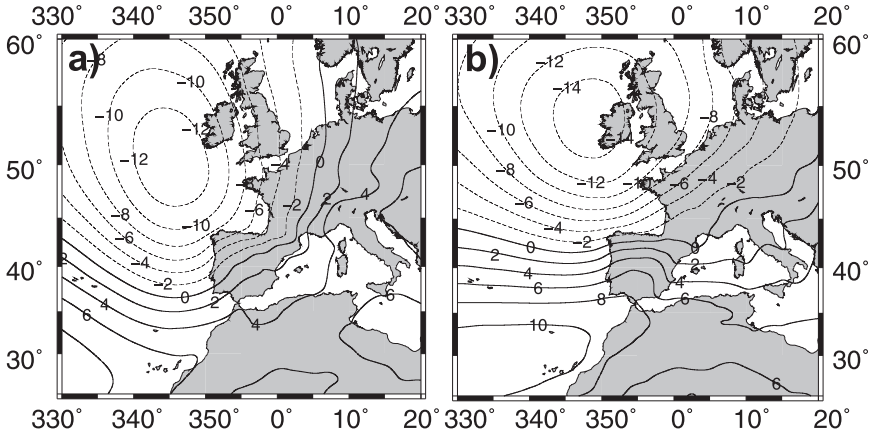

Fig. 9.3: Mean anomaly SLP fields (contour lines)for days from PP3 classified as WP5 and WP6 (b).

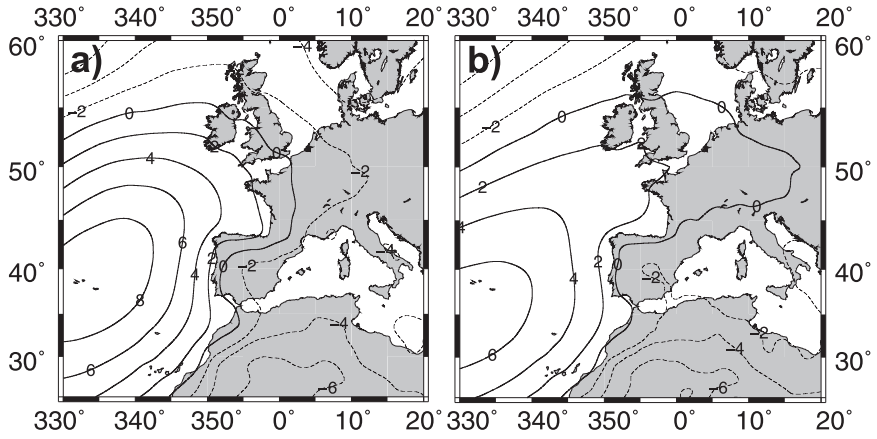

Fig. 9.4: Mean anomaly SLP fields (contour lines)for days from PP4 classified as WP1 (a) and WP2 (b).

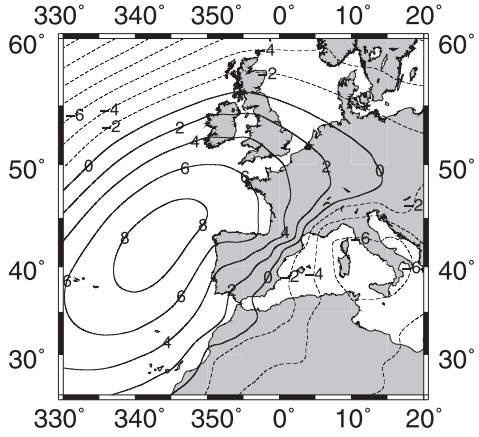

Fig. 9.5: Mean anomaly SLP fields (contour lines)for days from PP5 classified as WP1.

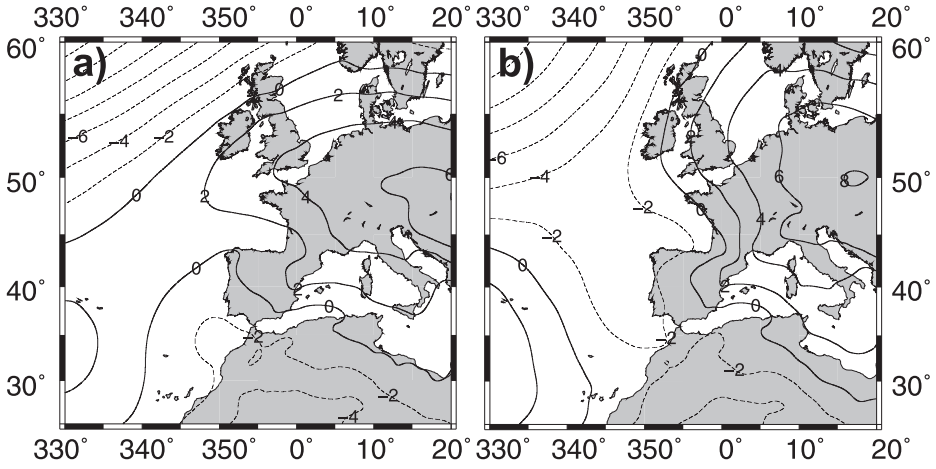

Fig. 9.6: Mean anomaly SLP fields (contour lines)for days from PP6 classified as WP4 (a) and WP5 (b).

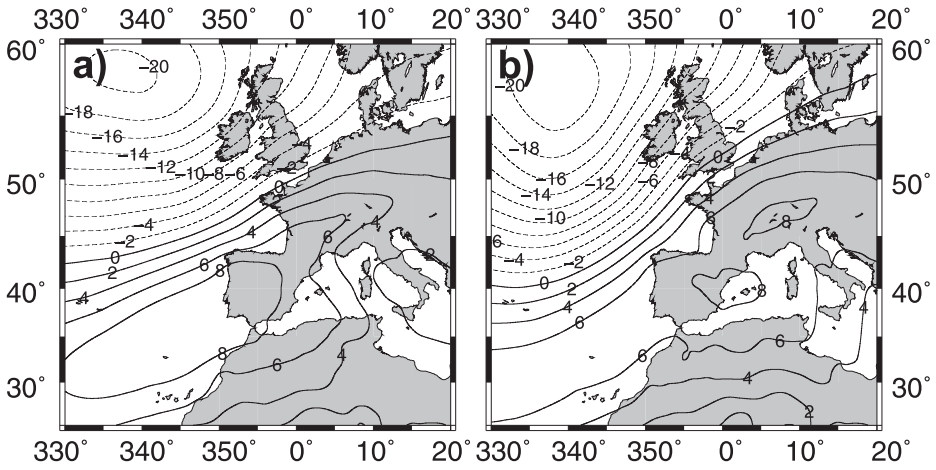

Fig. 9.7: Mean anomaly SLP fields (contour lines)for days from PP7 classified as WP3 (a) and WP4 (b).

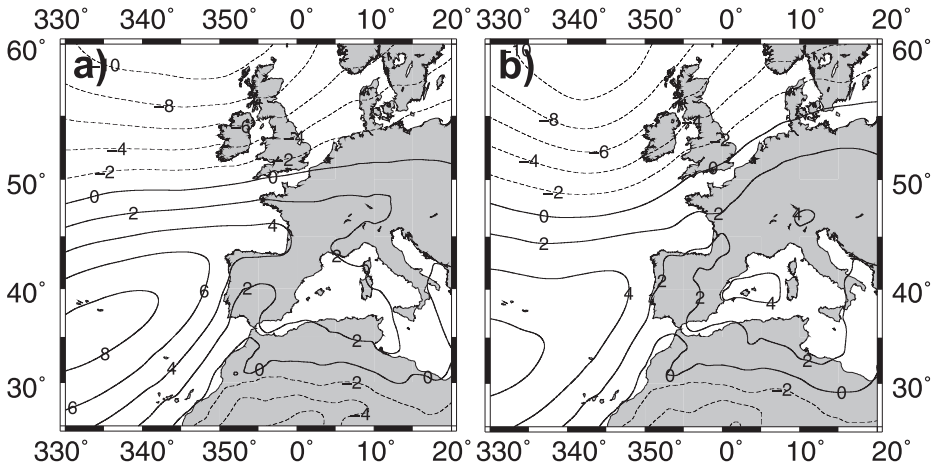

Fig. 9.8: Mean anomaly SLP fields (contour lines)for days from PP8 classified as WP2 (a) and WP4 (b).

Glossary